

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds  
(without alignments)  
4014.622 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFG.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR 79: \*  
1: Pirl: \*  
2: Pirl2: \*  
3: Pirl3: \*  
4: Pirl4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	5839	99.3	1153	1 RWHU1B	cell surface glyco
2	4447	75.6	1153	2 S00551	leukocyte surface
3	3456	58.8	1163	1 RWHU1C	cell surface glyco
4	1532.5	26.1	1170	2 S03308	cell surface glyco
5	1516.5	25.8	1163	2 S03206	lymphocyte fuction
6	1128	19.2	1179	2 A53213	integrin alpha-E c
7	1085.5	18.5	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45914	integrin alpha 2 s
9	1057	18.0	1178	2 A44142	integrin alpha-2 c
10	1054	17.9	1181	2 A33998	integrin alpha-1 c
11	1049	17.8	1180	2 A35854	lymphocyte-Peyer's
12	666	11.3	1039	2 A41131	integrin alpha-4 c
13	642	10.9	1038	2 S06046	integrin alpha-9 c
14	630	10.7	1035	2 I58409	integrin alpha cha
15	614.5	10.5	1041	2 T31437	cell surface glyco
16	579.5	9.9	1054	2 J07294	VLA-3 alpha subuni
17	572.5	9.7	1051	2 A35761	integrin alpha-5 c
18	567.5	9.7	1053	2 I55534	integrin alpha-V c
19	555.5	9.4	1053	2 S44250	integrin alpha-v c
20	551.5	9.4	1034	2 A36108	integrin alpha-6 c
21	539	9.2	1044	2 T10050	fibronectin recept
22	533.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	integrin alpha-5 c
24	531	9.0	1073	2 B36429	integrin alpha-6 c
25	530	9.0	1048	2 A27421	integrin alpha-5 c
26	529.5	9.0	1051	2 A40021	integrin alpha-3 c
27	524.5	8.9	1091	2 A41543	integrin alpha-6 c
28	514.5	8.8	1044	2 S16516	integrin alpha-8 c
29	497	8.5	1394	2 A29637	position-specific

#### ALIGNMENTS

##### RESULT 1

RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N: Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac-1; integrin alpha chain; neutrophil adherence receptor alphaM chain

C: Species: Homo sapiens (man)

C: Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004

C: Accession: A31108; A28915; A41600; A30892; A32218; A45526; A26091; 152867

R: Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A: Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CH

B.

A: Reference number: A31108; MUID: 88315033; PMID: 2457584

A: Accession: A31108

A: Molecule type: mRNA

A: Residues: 1-1153 <COR>

A: Cross-references: UNIPROT: P11215; GB: J03925; NID: g187284; PIDN: AAA59544.1; PID: g307148

A: Note: part of this sequence was confirmed by protein sequencing

R: Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A: Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M

A: Reference number: A28915; MUID: 88257215; PMID: 2454931

A: Accession: A28915

A: Molecule type: mRNA

A: Residues: 1-499,501-965, 'P', 967-1153 <ARN>

A: Cross-references: GB: M18044; GB: J03270; GB: M19664; GB: X07421; NID: g186935; PIDN: AAA594

A: Note: the authors translated the codon TAC for residue 1129 as Thr

A: Note: part of this sequence, including the amino end of the mature protein, was confir

R: Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A: Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg

A: Reference number: A41600; MUID: 92073318; PMID: 1683702

A: Accession: A41600

A: Molecule type: DNA

A: Residues: 1-9 <SHE>

A: Cross-references: GB: M76724; NID: g180018; PIDN: AAA58410.1; PID: g553215

R: Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A: Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi

A: Reference number: A94193; MUID: 88190151; PMID: 2833753

A: Accession: A30892

A: Molecule type: mRNA

A: Residues: 917-1042 <AR2>

R: Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A: Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor

A: Reference number: A32218; MUID: 89098893; PMID: 2563162

A: Accession: A32218

A: Molecule type: mRNA

A: Residues: 9-1153 <HIC>

A: Cross-references: GB: J04145; NID: g189068; PIDN: AAA59903.1; PID: g386975

integrin - fruit f  
integrin alpha-2b  
integrin alpha-7 c  
alpha-7 integrin -  
glycoprotein Iib -  
leukocyte adhesion  
integrin alpha cha  
F54F2.1 protein -  
integrin alpha v c  
hypothetical prote  
integrin alpha cha  
glycoprotein Iib -  
integrin alpha cha  
glycoprotein Iib -  
integrin alpha-1 -

30 496.5 8.4 1146 2 S40311  
31 496 8.4 1039 2 A34269  
32 491 8.4 1137 2 JC5950  
33 490 8.3 1135 2 I61186  
34 487.5 8.3 1037 2 A60163  
35 486 8.3 126 2 B30892  
36 470 8.0 1106 2 S38783  
37 466 7.9 1226 2 S44824  
38 458.5 7.8 1045 2 S60571  
39 444.5 7.6 1139 2 S28277  
40 418 7.1 1115 2 T09403  
41 413.5 7.0 1115 2 T09433  
42 391 6.7 784 2 I36316  
43 309 5.3 1086 2 T18523  
44 299 5.1 604 2 I36917  
45 279.5 4.8 272 2 A55348

Tue Nov 9 12:56:04 2004

A;Note: part of this sequence was confirmed by protein sequencing  
R;Pflaming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.  
J. Immunol. 150, 480-490, 1993  
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-1  
n during evolution.  
A;Reference number: A46526; MUID:93123748; PMID:8419480  
A;Accession: A46526  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-499,501-1153 <PLE>  
A;Cross-references: GB:S5227; NID:G263047; PIDN:AAB4821.1; PID:G263049  
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature  
R;Note: sequence extracted from NCBI backbone (NCBIP:121963)  
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.  
Biochim. Biophys. Acta 874, 368-371, 1986  
A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp  
A;Reference number: A90664; MUID:87076671; PMID:3539202  
A;Accession: A26091  
A;Molecule type: protein  
A;Residues: 17-31 <PIE>  
A;Experimental source: granulocytes  
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.  
Blood 79, 865-870, 1992  
A;Title: Characterization of the myeloid-specific CD11b promoter.  
A;Reference number: I52567; MUID:92144986; PMID:1346576  
A;Accession: I52567  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-9 <RES>  
A;Cross-references: GB:M84477; NID:G180184; PIDN:AA51960.1; PID:G553219  
A;Note: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
C;Genetics:  
A;Gene: GDB:ITGAM; CR3A  
A;Cross-references: GDB:120599; OMIM:120980  
A;Map position: 16p11.2-16p11.2  
A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag  
F;1-16/Domain: signal sequence #status predicted <Sig>  
F;17-1153/Product: cell surface glycoprotein CD11b #status predicted <Ext>  
F;17-1108/Domain: extracellular #status predicted <Ext>  
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA>  
F;465-473/Region: calcium/magnesium binding #status predicted  
F;530-538/Region: calcium/magnesium binding #status predicted  
F;593-601/Region: calcium/magnesium binding #status predicted  
F;1109-1134/Domain: transmembrane #status predicted <TM>  
F;1135-1153/Domain: intracellular #status predicted <INT>  
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.3%; Score 8939; DB 1; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCBPI 60  
Db 17 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCBPI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHTQTSNTYVKGCLCFLFGSNLRQPPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHTQTSNTYVKGCLCFLFGSNLRQPPQK 136  
Qy 121 FPEALRGCPQEDSDTAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTLESLMOYSEEP 180  
Db 137 FPEALRGCPQEDSDTAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTLESLMOYSEEP 196  
Qy 191 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240  
Db 197 RIHFTFKFQNNPNRSLVKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256  
Qy 241 TDGEKFGDPLGYEDVPELDRGVTRYVLGFGDAPRSKSRQELNTVASKPPRDHVFOAN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSKSRQELNTIASKPPRDHVFOAN 316

RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N;Alternate names: complement-3 receptor alpha chain

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C;Accession: S00551; 159078

R;Pyteila, R.

EMBO J. 7, 1371-1378, 1988

A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the

A;Reference number: S00551; MUID:88312584; PMID:3044779

A;Accession: S00551

A;Molecule type: DNA

A;Residues: 1-1153 <PVT>

A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G5298

A;Note: the authors translated the codon CAC for residue 569 as Gln

301 NPEALKTVONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNGPLLLSTVGSDWAG 360  
Db 317 NPEALKTVONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNGPLLLSTVGSDWAG 376  
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVILGAPYQHIGLVAMPR 420  
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILLRNRVQSLVILGAPYQHIGLVAMPR 436  
Qy 421 QNTGWESNANVKGTOIGAYFGASCSVDVDSNGSTDVLIGAPHYEYETRGQGVSCPL 480  
Db 437 QNTGWESNANVKGTOIGAYFGASCSVDVDSNGSTDVLIGAPHYEYETRGQGVSCPL 496  
Qy 481 PRGQARWQCDVAVLYGEQGPGRFGAALTVLGDVNGDKLTVDVAGAPEDNRGAVLYLF 540  
Db 497 PRGQARWQCDVAVLYGEQGPGRFGAALTVLGDVNGDKLTVDVAGAPEDNRGAVLYLF 556  
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQOSISGGQDLTMDGLVDLTGCAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFQOSISGGQDLTMDGLVDLTGCAQGHVLLRSQ 616  
Qy 601 PVLRVKAIMENPREVARNVFECDQVWKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660  
Db 617 PVLRVKAIMENPREVARNVFECDQVWKEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 736  
Qy 721 SLVGTPLSAPGNLRPVLAEADAQRLFTALPPEKNCNGDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAPGNLRPVLAEADAQRLFTALPPEKNCNGDNICODDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVTRNDGDSYRTQVTFPPDLDSYRKVSTLQNSQRSWRWLACESASSTEV 840  
Db 797 GPREFNVTVTRNDGDSYRTQVTFPPDLDSYRKVSTLQNSQRSWRWLACESASSTEV 856  
Qy 841 SGALKSTSCSINHIPIFENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNTEF 900  
Db 857 SGALKSTSCSINHIPIFENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNTEF 916  
Qy 901 QLELPVKYAVYVMVTVSHGVSTKYLNTFTASNTSRVMOHOYQVSNLQORSILPISLVFLVPV 960  
Db 917 QLELPVKYAVYVMVTVSHGVSTKYLNTFTASNTSRVMOHOYQVSNLQORSILPISLVFLVPV 976  
Qy 961 RLNOTVWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDIP 1020  
Db 977 RLNOTVWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVYVNCISIAVCQRIQCDIP 1036  
Qy 1021 FFGIOEEFNATLKGNLSDFDWYIKTSHNHLIIVSTABILFNDSTVFTLLPGQAFVRSOTET 1080  
Db 1037 FFGIOEEFNATLKGNLSDFDWYIKTSHNHLIIVSTABILFNDSTVFTLLPGQAFVRSOTET 1096  
Qy 1081 KVEPPEVNPPLPLIVGSSVGGLLALLALITAAALKYKGFKKQYKDMWSEGGPGEAPQ 1137  
Db 1097 KVEPPEVNPPLPLIVGSSVGGLLALLALITAAALKYKGFKKQYKDMWSEGGPGEAPQ 1153



R:Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986

A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement receptor type 1

A:Reference number: I59078; MUID:86287312; PMID:2942940

A:Accession: I59078

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 11-44 <RES>

A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193

C:Genetics:

A:Gene: Mac-1

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homologous protein

F:Keywords: cell adhesion; glycoprotein; transmembrane protein

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental

F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>

F:1106-1129/Domain: transmembrane #status predicted <TMW>

Query Match 75.6%; Score 4447; DB 2; Length 1153;  
Best Local Similarity 73.5%; Pred. No. 9.6e-303;  
Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFOENARGGQSVOVLGGSRVVVGAPOEIVAAORGSLYQCDYSTGSCEPI 60  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 17 FNLDEHPMTFOENAKGGQNVVLGGTSVVVAAPQEAKAVNGTGALYQCDYSTSRCHPI 76  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 61 RLQVPPEAVNMSSGLSLSAATSPPOLLAGCPTVHQTCSENTYVKGLCFEGSNLRQPOK 120  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 77 PLQVPPPEAVNMSSGLSLSAVSTPQQLLACGPTVHQCKENTYVNGLCYLEGSMLLRPQQ 136  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 121 FPPEARLGCQPQSDIAFLIDSGSIIPHDFRMKELVSTIMEQLKKSKTLFSLMQSYEEF 180  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 137 FPEARLECPQOESDITVFLIDSGSINNIIDFQMKKEPVSTVMGEQFKSKTLFSLMQSYDEP 196  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 181 RIHFTPEFQNPNRPSLIKITOLLGRTHRTATGRCVVRELFNITNGARKNAFKILFL 240  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 197 RIHFTFNDFKRNPERSHVPIKQNGRTKTASGIRKVVRRELPHKTNAGARENAKILVI 256  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 241 TDGEKFDPDLGYEDVIDELDREGVIRYVLGFGDAPERSEKSROELNVTASKPRPDHVFQAN 300  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 257 TDGEKFDPDLGYDVIPENDRAGVIRYVLGVGNAFNKPOSRRLEDLTIASPAGEHFVQVD 316  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 301 NFPAIKTVQNLREKIPIAEIGTOTGSSSSFHEMSQEGFSAAITSNGPLLSTVGSVDWAG 360  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 317 NFPAINTIQLOEKIFAIEGTQTGSTSFHEMSQEGFSAITSNGLPGSVGSFDWAG 376  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 361 GVPLYTSKEKSTFTNTRVDSMDNDAYLGXAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 420  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 377 GAFLYTISKDKVTINTTRVDSMDNDAYLGYSASAVILNRVQSLVLGAPRYQHIGLVAMFR 436  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 421 QNTGMVESNANRGTOIGAYFGALSVDVDSNGSTDVLVIGAPHYEOTRGQVSVCPL 480  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 437 ENFGTWEPHTSIKGSQIGSYFGALSVDMDADGNTNILLI GAPHYEKTGRGQVSVCPL 496  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 481 PRGORARWCDAVLVEQGQOPWRGFGAALTVLGVDSNGDKLTDVAITGAPCEEDNRGAIVLF 540  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 497 PRG-RARWOCEALLHGDDQGHFWGRFGAALTVLGVDSNGDKLTDVAIGAPEGQENQAVIF 555  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 541 HGTSGSGISPSHSORTAGSKSLPRLOYGOSLGGDLTMDGLVDLTVGAQSHVLLLRSQ 600  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 556 YGASTASLASASHRIIGHAFSPGLQYFGOSLGGDLTMDGLMDLAYGAQHLLLLRAQ 615  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 601 PVLRVKAIMEFPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRLEGDIQSTVT 660  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 616 PVRLREATMEFSPKVARSVFAQEQVILKNKDAGEVRVCLRVKNTKDRLEGDIQSTVT 675  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 661 YDALDLSGRPHSAVNETHKNSTRQTVLGLTQTCTETKLQLPNCIEDPVSPIVLRNF 720  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 676 YDALDLPVRSIRAFPEDTKNTRRRRTQVFLGMQKCTELLILPDCDVDSDSPIILRN 735  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 721 SILVGTPLSAFNGNLRPVLAEDAQRILFTALPFPEKNCNDNTICODDLSTIFSMLSCLVYG 780  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
QY 736 TLVGEPLRSFGNLRPVLAMDARFFTAFFPEKNCNDNSICODDLSTIMSAMGDTLVVG 795  
Db |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

Query Match 58.8%; Score 3456; DB 1; Length 1163;  
Best Local Similarity 60.6%; Pred. No. 2.5e-233;  
Matches 684; Conservative 143; Mismatches 296; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQBNARAGFGSVVQLOGSRVVVGAPOEIVAAANRGSLYQCDYSTGSCPEI 60  
DB 20 FNLDTTEELTAFRVDSAGFSDVQVYANWSVVVGAPOKITAANQTGGLYQCGYSTGACEPI 79  
QY 61 RLQVPVAVNMVGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFGLSGNLRLQPOPK 120  
DB 80 GLQVPVAVNMVGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFGLSGNLRLQPOPK 137  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQKXSKTULFSLMQSSEEP 180  
DB 138 LPVSRQCEPQEQDIFVLIDGSGSISSRNFAFMNFVRAVISQFQRPSTQFSLMQSSEEP 197  
QY 181 RIHFTKQFQNNPNRSLIKPITQQLQRTATATGLKRVVRELENTINGARKNAFKILFLL 240  
DB 198 QTHFTFEPRRTSNPLSLASVHQLQFTYTATAIQNVHRLHASTYGAARDATKILIVI 257  
QY 241 TDCEKFGDPIGYEDVIEPDLREGVIRVYVGLGDFRSEKSRQELNIVASKPPRDHVFQAN 300  
DB 258 TDGKKGESLDYKDVIPWADAAGIIRYVAGLAFQNNNSWKNELNDIASKESQEHIPKVE 317  
QY 301 NFEALKTVQNLREKIFAIEGTQGTSSSFEHMSQGFSAATISNGPLSTVGSYDWAG 360  
DB 318 DFDALKDIQNLKEKIFAIEGTQGTSSSFEHMSQGFSAATISNGPLSTVGSYDWAG 377  
QY 361 GVFLYTKBKSTFNTWTRVDSMDNDVILGYAAAIILNRVQSLVGLGAPRYOHIGLVAMER 420  
DB 378 GARLYPPNMSPTFNMQSDVMDRSDVGLSTELALWKGVSQVGLGAPRYOHIGLVAMER 437  
QY 421 QNTGWSNANVKGTOIGAYFGASLCSVDVDSNGSTDLVILGAPHYBYEYQTRGGQVSVCP 480  
DB 438 QVSRQWRKAEVTGQIGSYFGASLCSVDVDSNGSTDLVILGAPHYBYEYQTRGGQVSVCP 497  
QY 481 PRGQARQCCDAVLYGEGQPGWFGAALTIVLGVNVDGKLTVAIGAPGEEDNRGAVLYF 540  
DB 498 PRGWR-RWDCDAVLYGEGQPGWFGAALTIVLGVNVDGKLTVAIGAPGEEDNRGAVLYF 556  
QY 541 HGTSGSGISPHSQRIAGSKSLPRLOYGQSLSGGQDITMGLVDTLVAQGHVLLRSQ 600  
DB 557 HGVLPSPISPHSQRIAGSKSLPRLOYGQSLSGGQDITMGLVDTLVAQGHVLLRSQ 616  
QY 601 PVLRVKATMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVYT 660  
DB 617 PVLWVGSMQFIPAEIPSAFECRQVSEQTLVQSNICLYDKSKNLLGSRDLQSSVT 676  
QY 661 YDLALDSGRPHSRVAFNETKNSRRQTOVLGHTQTCETLKLQLPNCIEDPVSPVIRLNF 720  
DB 677 LDALDPGRLSPRATFQETKNSRSLRVRLVGLKAHCENFNLLPSCVEDSVTFITRLNF 736  
QY 721 SLVGTPLSAGNLRVLAEDAQLFTALPPEKNGNDNICODLSITFESMELDCLVVG 780  
DB 737 TLVGPPLAFNRLMLAADAORYFASLPPEKNGNDNICODLSITFESMELDCLVVG 796  
QY 781 GPREFNVTVTVNDGEDSVYRTQVTFPPFLDLRYKSVSTLQNRQSRQSWKLACASSTEV 840  
DB 797 SNLELNAEVMVNDGEDSVYRTQVTFPPFLDLRYKSVSTLQNRQSRQSWKLACASSTEV 854  
QY 841 SGALKSTGCSINHPPEPENSEVTFNITFDVDSKASLGNKLLKANTYSENMPRTNKTFF 900  
DB 855 SGTWSTSCRIINHLIPRGAQITFLATFDVSPKAVLGDRLLLTANVSSNNTPRTSKTFF 914  
QY 901 QLELPVKYAVYVTVSHGVSTKYLNFAS-ENTSRVQHQYQVSNLQSRSLPSLFLVLP 959  
DB 915 QLELPVKYAVYVTVSHGVSTKYLNFAS-ENTSRVQHQYQVSNLQSRSLPSLFLVLP 974  
QY 960 VRLNQTVWDRPQVTSSENLSTCHTKERLPSHDSFLAELRKAPVNVNCSTAVCQRQCDI 1019  
DB 975 VELNQAEVMDVEVSHQPNFSLRCSSEKIAAPPASDFLAHTQKNPVLDCSAGCLRFRCDV 1034

RESULT 4

S03308  
cell surface glycoprotein CD11a precursor - human  
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function C;Species: Homo sapiens (man)  
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text\_change 09-Jul-2004  
C;Accession: S03308; A47458; A47565; A48759; S36044  
J;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.  
J;Cell Biol. 108, 703-712, 1989  
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit A;Reference number: S03308; MUID:89139587; PMID:2537322  
A;Accession: S03308  
A;Molecule type: mRNA  
A;Residues: 1-1170 <LAR>  
A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UBC8; EMBL:Y00796; NID:931421; PIDN:CAA68 A;Note: part of this sequence was confirmed by protein sequencing  
R;Cornwell, R.D.; Gollahan, K.A.; Hickstein, D.D.  
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993  
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pr A;Reference number: A47458; MUID:93248261; PMID:8097887  
A;Accession: A47458  
A;Molecule type: DNA  
A;Residues: 1-20 <COR>  
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBIP:130863)  
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993  
A;Title: Identification of cell-specific and developmentally regulated nuclear factors A;Reference number: A47565; MUID:93281759; PMID:8099450  
A;Accession: A47565  
A;Molecule type: DNA  
A;Residues: 1-20 <SHE>  
A;Cross-references: GB:M95609  
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.  
J. Biol. Chem. 268, 19305-19311, 1993  
A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene promoter. A;Reference number: A48759; MUID:93374910; PMID:8103515  
A;Accession: A48759  
A;Molecule type: DNA  
A;Residues: 1-20 <NUE>  
A;Cross-references: EMBL:Z22804; NID:9311405; PIDN:CAA80461.1; PID:9311406  
C;Genetics:  
A;Gene: GDB:ITGAL; CD11A  
A;Cross-references: GDB:119757; OMIM:153370  
A;Map position: 16p11.2-16p11.2  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; t F;1-25/Domain: signal sequence #status predicted <SIG>  
F;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted F;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.1%; Score 1532.5; DB 2; Length 1170;  
Best Local Similarity 34.2%; Pred. No. 1.3e-98;  
Matches 401; Conservative 209; Mismatches 463; Indels 101; Gaps 37;

QY 1 FNLDTENAMTFQ-ENARGFGSVVQLOGSRVVVGAPOEIVAAANRGSLYQCDYSTGSCPEI 58  
DB 26 YNLDVFGARSPFPFRAGRFGYRVLQV-GNGVIVGAPGE---GNSTGLYQCQSGTGCHL 81  
QY 59 FRLQVPEAVNMVGLSLAATTSPPOLLACGPTVHQTCSNTYVYKGLCFGLSGNLRL-- 115  
DB 82 FVTLR-GSNVTSKYLGMTLATDPTDGSILACDPLGSLRTCDQNTYLSGLCYLFFQNLQGP 140  
QY 116 -QOQKFPFALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEQKXSKTULFSLM 174

Db 141 LOGRGFGQECIKG-----NVDLVFLFDGSMSLQDDEFQKILDFMKDVMKKLSNTSYQFAAV 196  
QY 175 QYSEEFRIHFTKEFQNNPNRSLIKPTITQLLGRHTHTATGLRKVYRELFNITNGARKNAP 234  
Db 197 QFSTSYKTEFSDYVKKKDPDALLKXVHMLLNTFGAINYVATEVREBELGARPDAT 256  
QY 235 KILFLLTGEKFGDPLGYEDVPELDREGVIRYVVLGFGDAFRSEKSRQELNVTASKPRD 294  
Db 257 KVLIIITDGE--ATDSGNIDAAD-----IIRYIIGIKHKFTKESQETLHKFKASPA 309  
QY 295 HVFOQANFEALTKVONQIREKIPALEGTQTGSSSSFEHMSQEGFSAAITNGPLSTVG 354  
Db 310 FVKILDTPEKLDLFTTEQKKIYVLEGTQSFLNFMELSSSGISADLSRGAHVAVGAV 369  
QY 355 SYDMAGGVF-LYTSKEKSTFNINTRVDSMDNDAYLGYAAA-IILNRVQSLVIGAPRYQH 412  
Db 370 AKDWAGGFLDLKADLQDDTFTGNEPLTEPVAGYLVYVWLPDRQKTSLLASGAPRYQH 429  
QY 413 IGLVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDILVIGAPHYEQ 470  
Db 430 MGRVLLFQEPQGGHWSQVQTHGTQIGSYFGGELCGVDVDDQGETELLIGAPLFYGEQ 489  
QY 471 RGGGVSCPLPRGORARWQCDV--LYGEQOPWGRFGAALTVLGDVNGDKLTDVAIGAP 528  
Db 490 RGRVFYI-----QRRQLGFEVSELOQDPGYPLGRFECAITALTIDINGDLVDVAVGAP 544  
QY 529 GEEDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMQGLVOLT 588  
Db 545 LEE--QGAVYIFNGRHG-GLSPQPSQRIEQTQVLSGQWFGRSIHGVKDLDEGDLADVAV 601  
QY 589 GAQGHVLLRSQPVLRVKALMEFNPREVARNVFECNDQV--KGKEAGEVRVCLHVQKSTR 647  
Db 602 GAESQIVLSSRPVDMVTLMSFSPABIPVHEVECSYSTSNKMKEGVNIICQI-KSLY 660  
QY 648 DLRLREGIQSWVYDIALDSGRPHSRVAFNETKNSTRQTVLGLTQCTBLKLQLPNCI 707  
Db 661 PQF-QGRVANLYTTLQDGHRTERRGLFPGGRHELRRNIATV-TSMSCDTDFGHPFCV 718  
QY 708 EDPVSPVILRNFLS---VGTPLS--AFGN-----LRPLAEDAQRILTALFPFEKNCN 757  
Db 719 QDLISPINVSLNLSLWEEESTPRDQRAQGDIPPIILRPSLHSETWEI-----PFKNCGE 773  
QY 758 DNTQDDLSITPSFMSLDCVLVGGPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVS 817  
Db 774 DKKEANRVSPSPARSRALRTAFASLSVELSLNLEEDAYVWQLDHPPLGSLRKYE 833  
QY 818 TLQORSORSRLACES--ASSTEVSGALKSTCSINHPIFPENSEVETNITFDVDSKAS 875  
Db 834 ML---XPHSQIPVSCBELPEERLLSRAL---SCNVSSPIFKAGHSVALQMMFNTLVNS 887  
QY 876 LGNKLLKANVTSENN---MPTNKTETQLEPLPVAVYVMVTVTSHGVSTKYLNFASEN 931  
Db 888 WGDSEVELHANVTCNEDSDLLDENSATTI---IPILYPINILIQOEDSTLYVSFPKGP 944  
QY 932 TSEVMQHQYQV---SNLQORSIP-LSLVFLPVRLNQTVWDRPQVTFSENLSSTCHTK- 986  
Db 945 KIHQVKMYQVRIQPSIHDPNIPTEAVGVFPQPESEGIPTHQWSQVMEPPV--PCHYED 1002  
QY 987 -ERLPSHSD--FLAELRKAPVNCSTAVCQRIQCDDIPFGIOEBFNATLKGNSFDWYTK 1043  
Db 1003 LERLFDAAEPLPGALFRCPVV-----FRQELVQVIGITLVLVGEIE 1044  
QY 1044 TSHNHLIVSTAEILFNDSVFTLLPQGAFAFRSQETKTVEPEVFNPLPLIVGSSVGLL 1103  
Db 1045 AS-SMFLSCSLISISFNSSKHPLHGYSNASL-AQVMKVVDVYVEKQMLYLYVLSIGGILL 1102  
QY 1104 LLALITAAIYKLGFPKQYKQVMMSESG-GPPGAEP 1136  
Db 1103 LLLLIFIVLYKGVFKRNLKKEQWAGRGVPGNIP 1136

RESULT 5  
156126

lymphocyte fuction-associated molecule-1-alpha - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I56126  
J:Kaufmann, Y.; Tseng, E.; Springer, T.A.  
J:Immunol. 147, 369-374, 1991  
A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit  
A:Reference number: 156126; MUID:91268576; PMID:2051027  
A:Accession: I56126  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1163 <RES>  
A:Cross-references: UNIPROT:P24063; GB:M60778; NID:g198785; PIDN:AAA39426.1; PID:g19878  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
F;151-315/Domain: von Willebrand factor type A repeat homology <WVAL>

Query Match 25.8%; Score 1516.5; DB 2; Length 1163;  
Best Local Similarity 33.9%; Pred. No. 1.7e-97;  
Matches 396; Conservative 216; Mismatches 466; Indels 89; Gaps 36;  
QY 1 ENLDTENAMTFQENA-RGFGSVVVLQGSRVVVGAPQEIIVAANQRGSLYOCDIYSTGCEP 59  
Db 24 YNLDTRPTQSLAQAGRHFGYQLQIEDG-VVVGAPGE--GNTGGLYHCRTSSEFCQP 79  
QY 60 IRLQVPVEAVNMSLGLSLAATTSPQILLACGPTVHOTCSENTYVKGLCFLFGSNLRQOPQ 119  
Db 80 VSLH-GSNHTSKYLGTLATDAAKGSLACDPGLSRTCDQNTYLSGLCYLFPQSLGPM 138  
QY 120 KFEALRGCPQEDSDTAFIDGSGSIIPHDPRMKELVSTIMBQLKSKTLFSLMYSEE 179  
Db 139 QNRPAYQECMKGVLDVLFVFDGSQLDRKDFEILEFMKDVMKLSNTSYQFAAVQFSTD 198  
QY 180 FRHFTFKEF-QNNPNRSLIKPTITQLLGRHTHTATGLRKVYRELFNITNGARKNAPKILF 238  
Db 199 CRTEFTFLDYVKNKPPDVLVLSGQPMFLTNTFRAINVYVAVHVFKEESGARPDATKVLV 258  
QY 239 LITDGEKF--GDPLGYEDVPELDREGVIRYVVLGFGDAFRSEKSRQELNVTASKPRDHY 296  
Db 259 IITDGEASDKGNISAAHD-----ITRYIIGIKHFVSQVKTLHIFASEPEEFV 309  
QY 297 FQANNEALTKVONQIREKIPALEGTQTGSSSSFEHMSQEGFSAAITNGPLSTVGSY 356  
Db 310 KILDTPEKLDLFTDIQRIYAIETGTRQDLTSFNLMELSSSGISADLSKGAHVAVGAK 369  
QY 357 DWAGGVF-LYTSKEKSTFNINTRVDSMDNDAYLGYAAA-IILNRVQSLVIGAPRYQHIG 414  
Db 370 DWAGGFLDLREDLQGAITFVQGPPLTSDVRGGYLVYVAMTSSRPLLAAGAPRYQHV 429  
QY 415 LVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDILVIGAPHYEQTRG 472  
Db 430 QVLLFQAPAEAGRWMTQTKIEGTQIGSYFGGELCSVDLDQDGEABELLIGAPLFFGEOR 489  
QY 473 QGVSVCPILPRGORARWQCDVLYGQOPWGRFGAALTVLGDVNGDKLTDVAIGAPGED 532  
Db 490 GRVFTY---QRRQLGFEVSELOQDPGYPLGRFGAAITALTIDINGDLTDVAVGAPLEE- 545  
QY 533 NRGAVYLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMQGLVLDITVGAQ 592  
Db 546 -QGAVYIFNGRPG-GLSPQPSQRIQGAQVFFGIRWFGRSIHGVKDLGGDLADVAVGAG 603  
QY 593 HVLLRSQPVLRVKALMEFNPREVARNVFECNDQVVKGEAG-EVRVCLHVQKSTRDLR 651  
Db 604 RWWLSSRPVVDVVTLSFSPEEIPVHEVECSYSAREEQKGVKLUKAFRIKPLFPQ--F 661  
QY 652 EGQTSVVYDIALDSGRPHSRVAFNETKNSTRQTVLGLTQCTETLKLQLPNCIEDPV 711  
Db 662 QGRLLANLSTQLDGHMRSRGLFPDGSHELSGNTSITP-DKSCLDHFFHFPICIQDLI 720  
QY 712 SPIVLRNFSLV---GTPLSAFGN-LRPVLAEDARLFTALFPFEKNCNNDNICDDLSI 767  
Db 721 SPINVSFNLSLEBEGTFRDQGRAMQPILRPSHTV-TKEIPFEKNGCKGKEANLTL 779

```

QY 768 TFSFMSLDCLVGGP-----REENVTVTVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQN 821
D 780 SSPARS-----GFLRLMSASLAVEWTLNSGEDAYWRLDLPFRGLSFFRKVEMLQ- 831
QY 822 QRSQRWRLACESASSTEVSGAL-KSTSCSINHIPIPEENSEVTFNITFDVDSKASLGNKL 880
D 832 --PHSRMPVSECEL--TEGSSLLTKUKCNVSPFIKAGEVSLQWNTLLNSMEDFV 887
QY 881 LLKANVTSEN-NMPTNKTEFOLELPVKYAVYVVTSHGVSTKYLNFTASENTSRVMOHQ 939
D 888 ELNGTVHCENENSLQEDNSAATHIPVLYPNVILTKQENSTLYISFTPKGPKTQOVQHV 947
QY 940 YQVSNLQORSLPISLVLNPNVRLNQTVIWRPQ-----VTFSENLS-----TCHTKE-RLP 990
D 948 YQV-----RIQPSAYDHNMT-LEALGVVPRPHSEDLTITVWSQTDPLVTVCHSEDKRP 1001
QY 991 SHSDFLAELRKAPVNVCSIAVCORIQDIPFGIOEBFNATLKNLSFDMWIKTSHNELL 1050
D 1002 SSE---AEQPCLPV-----QRCPIVF---RWELLQVTVGVLSKEIKAS-STLS 1046
QY 1051 IVSTAELFNDSVFTLLFGQAFYRSQTEKVPFEPVNPPLIIVGSGVGLLLALITA 1110
D 1047 LCSSLVSFNSSKHFHLYGSKA-SEAQVLVKVDLHEKEMLHVTVLSGIGLVLLFLIFL 1105
QY 1111 ALYKLGFPKROYKMM-SEGPPGAEP 1136
D 1106 ALYKVGFPKRLKMEADGGVPNGSP 1132

RESULT 6
A:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C:Accession: A53213
J:Shaw, S.K.; Cepak, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
R: Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit.
A:Reference number: A53213; PMID:94164962; PMID:8119947
A:Accession: A53213
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: UNIPROT:P38570; GB:I25851; NID:G457244; PID:G457245
C:Genetics:
A:Gene: GDB:ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo
F:199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.2%; Score 1128; DB 2; Length 1179;
Best Local Similarity 28.7%; Pred. No. 2.9e-70;
Matches 338; Conservative 213; Mismatches 454; Indels 172; Gaps 39;

QY 45 GSLYQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSIAATSPQQLACGPTVHQ 95
D 65 GPLHRCSLVQDEILCHPVEHVPFKGRHGVTVVRSHHGVLICI-----QVLVRRP--HS 117
QY 96 TCSEYTVYKGLCFGLFGLNLRQVQO-----119
D 118 LSESLT---GTCLLIGFDLRPQAQANFFDLNLLDPADVTDGDCYCNKEGGEDDVNTA 174
QY 120 KPEALRGCPQED-----SDIAFLIDGSGSIIPHDFRMRKELVSTIMSQL-- 164
D 175 RQRALEKEEEDKEEEDDEEAEAGTETAILIDGSGSDPPDFQAKDFISNMNRNFEY 234
QY 165 KSKTTLFSLMOYSEEFRIHTFKFQFNPNRSLIKPIITQLLGRTHATGLRKVVRELEN 224
D 235 KCFCNPFALVQYGVGVITQTEFLRDSQDVMASTARVONITQVGSVTKTASQHVLSIFT 294
QY 225 ITNGARKNAKFLILLTDGKFGDPLGVEDVPELOREGVIRYVVLGFDGAFRSEKROEL 284

```

```

D 295 SHGSRKSKVMMVVLTDGIFEDPLNLTVINSKMGQVVERPAIGVGEFFKSARTAREL 354
QY 285 NTVASPRPDHVPQANNFEALKTQVQNLREKIFAIEGTQTGTSSSSSPEHMSQGFSAAIT 344
D 355 NLITASPDDETHAFKVTNYMALDGLLSKLRNIIISMEST---VGDALHYQLAQIGFSAQIL 411
QY 345 SNGP-LLSTVGSYDAGWGVFLY--TSKEKSTFINNTRVDSMDNA-----YLYGAAAILRN 398
D 412 DERQVLLGAVGDFWSGALLYDTRSGRFLNQTAATAAADAQAYSYLYGAVAVLHKT 471
QY 399 RVQSLVLGAPRYOHIGLVAMFR--QNTGMWESNANV-KGTQIGAYFGFASLCSVDVDSNGST 456
D 472 CSLSYVAGAPQYKHG--AVFELQKEGREASFUPVLEGEQMGVSFGSELCPVDIDMDGST 529
QY 457 DLVLIGAPHYYEOTRGQVSVCPPLPRQQRARWQCDVLYGEGQGPWRFGAALTVLGDVN 516
D 530 DFLVLAAPFVHVHGEGRVYVRLSE-QDGSFSLARILSGHPGFTNARFGFAMAAMDLS 588
QY 517 GDKLTDVAIGAP-----GEEDNR--GAVLPHGTGSGISPSHSQRIAGSKLSPLOYFGQ 570
D 589 QDKLTDVAIGAPLEGFGADGASFGSVIYNG-HWDGLSASPSQIRIRASTVAPGLQVFGM 647
QY 571 SLSGQDLTMDGLVDLTGAGQHVLLRLRSOPVLKRVKAIMFENPREVARNVPCNDQVVK 630
D 648 SMAGGFDISGDGLADITVGLQAVVFRSPVRLKVKVMAFTPSALP-----IGF 697
QY 631 KEAGEVRVCLHVQKSTRDLREGIQSVTVYDLALDSGRPHSRFAVFNETKNSRROTQVL 690
D 698 NGVNVRLCFEI--SSVTASSEGLREALNFTLVDVGKQRRRLQCSQSVRCLGCLREWS 756
QY 691 GLTQTCETLKLQLPN-----CIEDPVSPVILRLNLSLVTPLSAFGLNLPVLAEDARLFT 746
D 757 SGSQLCEDL-LIMPTEGELCEEDCFSNASVSVYQL-QTPREGQTDHPQFILDRTYEPFAI 814
QY 747 ALFPFKMCGNDNICODDLSITFSMSLDCLVGGSPREENVTVTVRNDEGDSYRTQVTF 806
D 815 FQLPYEKACKNKLFCVABLQLA-TTVSQOELVVGULTKELTLNINLTNSGSDSYMSTMALN 873
QY 807 FPLDLSYRKVSTLQNRQSRWRLACESASTEVSGALKSTSCSINHPIFENSEVTNI 866
D 874 YPRNLQ-----LKEMQKPPSPNICDDPQPV---ASVLMNCRIGHFVL-KRSSAHVSV 923
QY 867 TFDVDSKASLGNKLLKANVTSENN-----MPRTNKTEFO---LELPVKAVYVTVSHGV 919
D 924 VMOLEENAFPNRTADITVTVINSNRRSLANETHLQFRHGFVAVLSPKSIWYVNTGQGL 983
QY 920 S--TKYLNFTASENTRVWQHOYQVSNLQORSLPISLVLVPLVRLNQTVIMDRPQVTFSE 977
D 984 SHHKBFPLFVHGEN---LFGAEYQ-----LQICVTKLRLGLQVAAVKLRTQ 1028
QY 978 NLSSTCHTKERLPHSHSDFLAELRKAPVNVNCSIAVCORIQDIPFGIOEBFNATLKNLS 1037
D 1029 ASVTCTWQSERACAYSS-VQHVVEHWSVSCVIA-----SDKNVTVAAEITS 1073
QY 1038 PDWYIKTSHNLLIVST-----ABILENDSVFTLLPGQAFVRSQTEKVPFEPVNP 1091
D 1074 WD-----HSELLKDVTELQILGETISFNKSIYEGINAENH--RTKITVTVFLKDEKXHS 1125
QY 1092 PLIVGSSVGLLALLITAAVKLGFFKQYKDMSE 1128
D 1126 PIIIKSGVGLLVILVILVILFCGFFKRYQOOLNLE 1162

RESULT 7
A:Species: Homo sapiens (man)
C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A45226
R: Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J: Biol. Chem. 268, 2989-2996, 1993
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit

```

integrin alpha-1 chain - human (fragment)



370 VGFSAEYSPQNILMLGAVGAYDWSGTVVQKTPHGHILIFSQAPEQLQDRNHSSVLGYS 429  
392 AATILNRVQSLVGLGAPRYOHIGLVAMFRONTGMWESNANV-----KGTQIGAYFGASL 445  
430 VASISTGNSVHFVAGAPRANYTQIVLYSVN-----ENGNVTIVQSQRGDIQSYFGSVL 484  
446 CSVDVDSNGSTDVLVLGAPHYBQTR--GGOVSVCLPRQARARWQCDVLYGEQGPWG 503  
485 CAVDVNDKTTITDVLVGPAPMYMDLKKKEGRVLTITK-ILNWH--QFLEPGNGLENA 541  
504 RFGAALTIVLGDVNGDKLTVAIGAPGEDNRGAVLPHGTSGSGISPSHQRAGS--KL 561  
542 RFGSAIAALSDINMDGNDVIVGSPLENQNSGAVIYNGHEG-IRLRYSQKILGSDRAF 600  
562 SPRLOYFGQSLSGQDITMDGLDITVGAQGHVLLRSQPVLRVKAIMEPNPREVARNVF 621  
601 SSHLOYFGRSLDGYDLNGDSITDVSAGAFQVQVQLWSQSIADVSVDASFTPKKI--TL 658  
622 ECNDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVYDIALD---SGRPHSRAVEN 677  
659 NKNAEI-----KLKLCF-----SAKFRPTNQNNQVAIVNITIDEQFSRVSIRGLFK 707  
678 ETKNSTRTQTVLGLTOTCE--TLKLQLPNCIEDPVPSPVLRVLRNLSL--VGTPLSAFGL 733  
708 ENNERCLOKTMVSAQACRSBYIIHQEPS---DIISPLNLCMISLENPCT-----756  
734 RPVLAEADAQRLFTALFFPEKNCNGNDNICQDLSITP-----SFMSLDCLVWGGPREFNTV 789  
757 NPALAEYSETVKVPSIFPHKCGDDGVCISDLVNVQLPATQQQPFIVSNQNKRLTFSV 816  
790 TVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASST-EVSGALKSTS 848  
817 QLKNNKESAYNTEIIVDFSENLF-----ASWMPVDGTEVTQIASSQKSVT 864  
849 CSINHPPIENSEVTNFTEDVDSKASLGNKLLKANVTSENMPRTNKTPEFQLELPVKY 908  
865 CNVGPALKSKQCVTTFINDFNLQ-NLQONASISFRALSESQENMADNSVNLKSLLY 923  
909 AYVMVVTSHGVSTKYLNTASENTSRVMQHOYQVSNLQOR-----SLPISLVFLV 958  
924 DAEIHIT-RSTININFEVSLDGNVSVV-HSFE-DIGPKFISIKVTITGVSFVMSMA---976  
959 PVRLNQTIVDRPQVTFSEN---LSTCHTKE-----RLPSHSDPLAB- 998  
977 -----SVLIHTPQYTKDKNPLMYLTGVHTDQAGDISCEARINPLKIGOTSSSVSKSEN 1030  
999 LRKAPVWNCIAVCQRIQCDIPFGIOEEFNATLGNLSFDWYIKTSHNHLIVSTAET- 1057  
1031 FRHIKELNCRITASCSNIMCWLRDLQVKGEYFLNVSTRWNGTFAASTFQVLTAAAEID 1090  
1058 LFNDVSFTL-----LPGOGAFVRSQTEKVEPFE-VPNPLPLIVGSSVGGLLLLALITA 1110  
1091 TYNPQIVIEENTVITP-----LTIKPKHEKVEVPTGVIVGSIAGILLLLALVA 1140  
1111 ALYKLGFFKRYQKDM 1125  
1141 ILMKLGFFKRYEKM 1155

RESULT 9  
S44142  
VLA-2 protein homolog - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S44142  
R;Edelman, J.M.; Chan, B.M.; Uniaval, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze  
submitted to the EMBL Data Library, January 1994  
A;Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not  
A;Reference number: S44142  
A;Accession: S44142  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1178 <EDE>

A;Crcss-references: UNIPROT:Q62469; EMBL:Z29987; NID:9473098; PIDN:CAA82877.1; PID:94730  
F;169-344/Domain: von Willebrand factor type A repeat homology <WMA2>  
Query Match 18.0%; Score 1057; DB 2; Length 1178;  
Best Local Similarity 27.8%; Pred. No. 2.7e-65;  
Matches 341; Conservative 209; Mismatches 487; Indels 188; Gaps 45;  
QY 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGSRVVVGAPQEIIVAANQSGSLYOC--DYST 54  
Db 27 YNVGLFGAKIFSQSPSEFGYVQQLTNPQGNLLVGSWSPGPNRMGVDVYKPCVDLPT 86  
QY 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATSPQILLACGPTVHQTCSENTYVKGCL 107  
Db 87 ATCEKLNQNSASISNVTEIKTNMSLGLITRNPGTGGFTLTCGFLWAHQCNQNYATGIC 146  
QY 108 FLGSLNRLOQPO---KPEALRGCPQEDSDIAFLIDGSGSIIIPHDFFRMKELVSTIMEQL 164  
Db 147 ----SDVSPDFQLTSPSPAVQACPSL-VDVVVVCDSESNLYP--WEAVKNFLVKFTGL 199  
QY 165 K--KSKTLFSLIMQYSEEFRIHFTKFEQNNPNRSLIKPTQLLQ-RTHATGLRKVVRE 221  
Db 200 DIGPKTQVALIQVANEPRIFNLNDFETKEDMVQATSETRQHGGLTNTFRAIEFARDY 259  
QY 222 LFNITNGARKNAFKILFLITDGEKFGDPLGYEDVIPELDRGVIRY---VLGE--GDAFR 276  
Db 260 AYSQTSGRPGATKVMVWVTDGESH-DGSKLKTVIQCNDDDEILRFAGVAVLGYLNRNALD 318  
QY 277 SEKSRQELNTVASPPRDHVFQANNFALKTVQNLREKIFAIEGTOTGSSSPHEMSQ 336  
Db 319 TKNLKEIKALASTPTERYFENVADAALLEKAGTLGBOIFSIETGVQ-GDNFQMEWAQ 377  
QY 337 EGFA--AITSNGPLLSVGSYDWAGGVFLVTSKSTFIINMT--RVSDMN-DAYLGYA 391  
Db 378 VGFSAIYAPQNDILMLGAVGAFDWSGLVQETSHKPVIFPKQAFQVQLQDRNHSSFLGYS 437  
QY 392 AATILNRVQSLVGLGAPRYOHIGLVAMFRONTGMWESNANV---KGTQIGAYFGASLCS 447  
Db 438 VAAIATEDGVHFVAGAPRANYTQIVLYSVNK---QGNVTIVQSHRGQIGISYFGSVLCS 494  
QY 448 VDVVNGSTDLVLGAPHYBQTR--GGOVSVCLPRQARARWQCDVLYGEQGPWGPRF 505  
Db 495 VDVVDKTTITDVLVGPAPMYMDLKKKEGRVLTITKILNQHQ---FLEGPEGTGNARF 551  
QY 506 GAALTIVLGDVNGDKLTVAIGAPGEDNRGAVLPHGTSGSGISPSHQRAGSGLSPR- 564  
Db 552 GSAIAALSDINMDGNDVIVGSPVENSGAVIYNGHQT-IRTKYSQKILGSGNGAFER 610  
QY 565 -LOYFGQSLSGQDITMDGLDITVGAQGHVLLRSQPVLRVKAIMEPNPREVARNVFEC 623  
Db 611 HLQFFGRSLDGYDLNGDSITDVSIGALGQVQLWSQSIADVAIEALFTP-----660  
QY 624 NDQVVKGEAGVRVCLHVQKSTRDLREGQIQSVVYDIALD---SGRPHSRAVENET 679  
Db 661 -DKITLLNKDAKITLKLCPRAEFRRAGQNNQV--ALLFNMTLDDADGSHSRVTSRGVREN 717  
QY 680 KNSTRRTQVGLTQTCET--IKLQLPNCIEDPVPSPVLRVLRNLSLVTGTFSAFGLNRPVL 737  
Db 718 SERFLQKNVWVNEVQKSEHHISIQKPS---DVNPLDLRVDISLENPGTS-----PAL 768  
QY 738 AEDAQRLFTALFPFEKNCNGNDNICQDLSI-----TFSPMSLDCLVWGGPREFNVMT 788  
Db 769 EAYSETVKVPSIFPYKEGSDGICISDLILDVQQLPAIQTSF-----IVSNQNKRLTFS 823  
QY 789 VTVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASST-EVSGALKST 847  
Db 824 VILKNGESAYNTVILAESENLF-----AGSPMPVDGTEVTCEVGSQKSV 871  
QY 848 CSINHPPIENSEVTNFTEDVDSKASLGNKLLKANVTSENMPRTNKTPE--FQLELP 905  
Db 872 TCDVGVPAKSEQCVTTFINDFNLQ-NLQONAAINFQAFSESQ--ETNKADNSVSLTIP 928  
QY 906 VKYAVYVMTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQOR-----SLPISLV 955



```

Db 929 LLYDAELHLLT-RSTNINFEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSWA 984
QY 956 FLV-----PVLNQTVIWRDPQVTF-SENLS 980
Db 985 LVTHIPIQYTKENPLLYLTGLTQDQAGDISCTABINPLKLPHTA-----PVSFKENFR 1040
QY 981 STCHTKERLPKSHDLAEARKAPVNVCSIAVCQRIQCDDIPFFGIOEEFNATLKGNLSPDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMKAEYFINVTVVWRT 1080
QY 1041 YIKTSHNHLIVSTAAILFNDSVFTLLPGQAFVRSQTETKVEPPEVNPPLIVGSSVG 1100
Db 1081 FAASTFQTVQTAAREIDTHNPQFVBIENAVTIPLMIMKTEKAEVPT--GVIIGSIIA 1138
QY 1101 GLLLLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLLLAMTAGLWKLGFQYKKM 1163

RESULT 10
A33998
Integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A;Reference number: A33998; MUID:9308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:PR07301; GB:X17033; NID:G33906; PIDN:CAA34894.1; PID:G33907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and GPIIb)
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutcher, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:L24121; NID:9400342; PIDN:AAA16619.2; PID:94583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VMA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match
Best Local Similarity 17.9%; Score 1054; DB 2; Length 1181;
Matches 328; Conservative 215; Mismatches 494; Indels 188; Gaps 44;

QY 1 FNLDTENAMTFQ-ENARFGGSVVQL---QGSRRVVVGVAPQEIIVAAVNRGSLYQC--DYST 54
Db 30 YNVGLPEAKIFSGPSSEQFGVAVQCFINPKGNWLLVGVSGFPPENRMDGVKCPVDLST 89
QY 55 GSCPEIRLQ-----VPVEAVNMSLGLSLAATSPPOLLAACGPTVHQTCSNTYVKGLC 107

```

```

Db 90 ATCKLNLQTSSTIPNVTEMKTNMSLGLILTRNMGTCGFLTCGLPWAQCCGQNYTTGVC 149
QY 108 FLFGSNLRQQQKPEALRGCPQEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLK-- 165
Db 150 SDISPDF-OLASASPSATQPCPSI-IDVVVCDSDNSIYPWD--AVKNFLEKFFVQGLDIG 205
QY 166 KSKTLFSLMOYSEEFRIHFTFKFQNNPNRSLKIPITQLLG-RTHATGATGRKVVRELFN 224
Db 206 PTKTQVGLIQVANNPRVVFNLTVYKTEEMIVATSQTSQYCGDLTNTFGAIQYARKAYS 265
QY 225 ITNGARKNAPKLELLTDGKFGDPLGYEDVIPEDREGVIRY---VLGF--GDAFRSEK 279
Db 266 AASGRRSATKVVVVVTDGESH-DGSMKKAVIDQCHDNILRFGIAVLGLNRLNALDTKN 324
QY 280 SRQELNTVASKPRDRHVFOANNFEALTVQNQLREKIFAETGTGTGSSSSPEHEMSQGGF 339
Db 325 LIKEIKAIASIPTRYFFNVSDAALLEKAGTIGEIFSTEGTVQG-GDNFQMEMSQVGF 383
QY 340 SAAITSNRP--LHSTVGSYDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGAAAAI 394
Db 384 SADYSSQNDILMLGAVGAPCWSGTIVQKTSHGHLIFPKQAFDQLQDRNHSVLYGVAA 443
QY 395 ILNRVQSLVILGAPRVOHIGLVAMFRONTGOWBSNANV-----KGTQIGAYGASLCSV 448
Db 444 ISTGESTHFVAGAPRANYTQGIIVLYSVN-----ENGNITVIQAHRGDQIGSYFGSVLCV 498
QY 449 DVDSNGSTDLVLICAPHYEQTR--GGOVSVCPLPRQARQWQCDVAVLYEQOQPGWRFG 506
Db 499 DVDKDIITDVLIVGAPMYMSDLKKEGRVYLFYTKKILGQHQ---FLEGEPIENTRFG 555
QY 507 AALTVLGVNGDKLITDVAIGAPGEEDNRGAVLYFHGTSGSGISFHSORTAGS--KLSPR 564
Db 556 SAIAALSDINMDGFNDVIVGSPLENQSGAVIYNGHQGT-ITKYSQKILGSDGAFRSH 614
QY 565 LQYFGOSLGSQGLTMDGLVDLTVGAQHVLRLRSQPVLRVKAIMEFNPREVARNVFECN 624
Db 615 LQYFGSLDGYGDGLNGDSITDVSIGAFGVQVQLWSQSIADVAIEASTPEKI--TLVKN 672
QY 625 DVVKGKEAGEVRVCLHVQKSTRDLRREGIQSVVTVYDLALD-----SGRPHSEAVNETK 690
Db 673 AQII-----LKLCE---SAKFRPKQNNQVAIVINITLDADGFSSTVSRGLFKENN 721
QY 681 NSTRQTVLGLTQTC--ETLKLQPLNCIEDVSPVILRLNFSILVGTPLSAFGLNRPVLA 738
Db 722 ERLCKNMVWVNAQSCPEHIIYIQEPS--DVVNSLDLRVDISLENPGTS-----PALE 772
QY 739 EDAQRLFTALPEKFNKCGNDICQDDLSITF---SFMSLDCLVVGPPREFNVTVTRND 794
Db 773 AYSETAKVFSIPFHKDCDGEDGLCISDLVLDVRQIPAAQEPPIVSNQNKRLTFSVTLKNK 832
QY 795 GEDSYRTQVTFPPLDLSVRKYSTLQNRQSRKSWRLACESAST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTEVTCQVAASQKSVACDVG 880
QY 854 PIPPESENVTNITFDVDSKASLGNKLLKANKVTSENNMPTNKTEFOLELPVKVAVMV 913
Db 881 PALKREQQVTFITNFDNLQ-NLQNAQSLSCALSESQENKADNLVNLKIPLLYDAEI- 938
QY 914 VTSCHVSTKYLNFTASENTSRVMOHQVQVSNLQOR-----SLPISLVFLV----- 958
Db 939 ---HLTRSTNINFEISDGNVPSIVHVSFEDVGPKFIFSLKVTGTGSPVSMATVIIHPQ 995
QY 959 -----PVRLNQTVIWRDPQVTF-SENLSSTCHTKER 988
Db 996 YTKENPLMYLTGVQTDKAGDISCNADINPLKIQGT-----SSVSFKSENFR---HTKE- 1047
QY 989 LPSHSDFLAELKAPVNVCSIAVCQRIQCDDIPFFGIOEEFNATLKGNLSFDWIYKTSNH 1048
Db 1048 -----LNCRTASCNVTCWLKDVHMKGEYFVNTVTRNWTFFASSFTQT 1091
QY 1049 LLIVSTAEI-LFNDVSFTLLPGQAFVRSQTETKVEPPEVNP-----LP--LIVGSSVG 1100

```



Tue Nov 9 12:56:04 2004

Db 1092 VOLTAAAEINTNPELYVI-----EDNTVPIPLMKPDEKAEVPTGVIIISITA 1141

Qy 1101 GILLALLALITAAALYKLGFFKRYQKDM 1125

Db 1142 GILLALLALVAILWKLGFKKRYEKM 1166

RESULT 11

A35854

integrin alpha-1 chain precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 23-Oct-1990 #sequence\_revision 13-Sep-1991 #text\_change 09-Jul-2004

C;Accession: A35854; S11243

R;Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990

A;Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin A

A;Reference number: A35854; MUID:90338125; PMID:2380249

A;Accession: A35854

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1180 <IGN>

A;Cross-references: UNIPROT:P18614; GB:X52140; NID:G56493; PIDN:CAA36384.1; PID:G56494

C;Keywords: cell adhesion; cytoskeleton; transmembrane protein

F;170-345/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 17.8%; Score 1049; DB 2; Length 1180;

Best Local Similarity 27.3%; Pred. No. 9.9e-65;

Matches 344; Conservative 200; Mismatches 476; Indels 240; Gaps 48;

Qy 1 FNLDTEAMTFQENARG-EGOSVVLQ---QGSRRVVGAPQEIIVAANQSGSLYQCDYSTGS 56

Db 29 FNDVKNKMSGSPGVDMFGVTVOQYENEEKWVLIGSLVPLVGQPKARTGVYKPCVGRER 88

Qy 57 CEP-IRLQVPVEA-----VNMSLGLSLAATTPSPQLLACGTFVHQTCSNTYVKG 106

Db 89 AMPVCUKLDFVNTSPNTEIKENMTFGSTL-VINPENGFLACGLPLAYRCGHLHYTTGI 147

Qy 107 CFLFGSNLRQOPQKPEALRGCPQSDIAFLIDGSGIIPHDFRMKELVSTIMEQLKK 166

Db 148 CSDVSPTFQVNVSEFAP--VQECSTQ-LDIVIVLDGNSIYP-----WESVIAFLNDLKKR 199

Qy 167 -----SKTLFSIMYSEFRHFTEKFEONNPNRSLIKPTQLLG-RTHATGLRKVVR 220

Db 200 MDIGPKQTVGVIVQGENVTHEFNLNKYSSTEEVLVAANKRGQGLQTMALGIDTARK 259

Qy 221 ELFNITNGARKNAFKILFLTLTDERKFGDPLGYEDVIPELDRGVIRY---VLFGDA--F 275

Db 260 EAFTEARGARRGVKKVMIVTDGESH-DNYRLKQVIQDCEDENIQRFSTAILGHYNRGNL 318

Qy 276 RSEKSRQELNTVASKPPRDHVFQANNFEALKTQVQLREKIFAIEGTQTGSSSFEHEMS 335

Db 319 STEKFEETKSTASEFTKHFNFVSDDELALVITVIALGERIFALEATADQSAASEMENS 378

Qy 336 QGFGAAITNSGPLASTVGSYDAGGVFLYTSKESKSTFINMT--RVDSMDND---AYLGY 390

Db 379 QTFGSAHVSQDMWMLGAVGAYDNGTVMQKANQWVPHNTFTQTEPAKNCEPLASYLGY 438

Qy 391 AAAAILNRVQSLVLGAPYQHIGLVAMFRONTGMWESNANVKGTOIGAYFGASLCSDV 450

Db 439 TVNSATIPGDVLYIAGQPRYNHTGVQVLYKMDGNGINILQTLGGEQIGSYFGSVLTIDI 498

Qy 451 DNGSTDLVLIGAPHY-----YEOTR-GGQVSVCPPLRGORARWQCDAVL 494

Db 499 DKDSYTDLLVGLAPMTYMTGTEKEEGKGVYVAVNQTTFEYQMSLEPIRQTCSSLKNSCT 558

Qy 495 YGFGQGPWG-RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYFHGTSGSISPSHS 553

Db 559 KENKNEPCGARFGTATAANKDLNVDGNDVIGAPLEDHAGAVIYHG-SGKTIREAVA 617

Qy 554 QRIAGSKSLPRLQFYQSGGQDLTWDGLVDLTVGAQGHVLLRSQPLVRKVAIMEFNP 613

Db 618 QRIPSGGUGTKLKFPGQSIHGEMDLNGDLGTLDTVTIGLGAALFWARDVAVVKVMTNFP 677

Qy 614 REVARNVPECNDQVQVKGKEAG--EVRVCLHVQ-KSTRDLREGTQSQVVYDIALDLSGRP 670

Db 678 NKVNIQKNCR---VEGKETVCINATMCPHVKLSKESDIYADLQ---YRVTLSDLRQ 730

Qy 671 HSRVAFNET-----KNSTRRTQVLGLTQTCETLKLQLPNCI-----EDPVSPIV 715

Db 731 ISRSFSGTQERKIQRNITVRESE-----CIRHSFYMLDKHDFQDSVR 773

Qy 716 LRLNFSLVGTPLSAFCNLRPVLAEADAQRLFTALFFPEKNCNDNI CODLSITFSFMSLD 775

Db 774 VTLDFELT-DPENG-----PVLDDALPNSVHEHIPAKDCGNKERCISDLTINVSTTEKS 827

Qy 776 CLVVGSGPRE-FNVTVTRNDGEDSYTQTFPPDLDSYRKVSTLQNRQSRQSRWELACES 834

Db 828 LLIVKSQHDKNFVSLVAKNGDSAYNTRTVVQHSPLNIFSGIBEIOKD-----SCS 879

Qy 835 ASSTEVSGALKSTSCINHIPIFENSEVTFNITFDVDSKASLGNKLL-LKANVTSENMP 893

Db 880 N-----ONITCRVGFPLRAGETVIFKILFQFNWTHLSENAIIHLSATSDSEEPLE 930

Qy 894 RTNKTEFOLELPVKYAV---YMWVTSHGVST-----KYLAFNTASENTSRVMQHOYQV 942

Db 931 SLMDNEVNISIPVKYEVGLQFYSSASEHHISVAANETIPEFINST--EDIGNEINVFYTI 988

Qy 943 SNLQORSPLP--ISLVP-----LVPVRLNQTVWD-----RP----- 971

Db 989 RKGHFMPELQLSISFPNLTAGYPVLYPIG-----WSSDNVNCRPSLEDPFGINS 1042

Qy 972 --QVTFPS-----ENLSSSTCHTKERLPSHSDFLAELKAPVNVNCISIAVCORIQCDI 1019

Db 1043 GKMTITSKSEVLKRGTTIQDCSSTC-----GVATITCSLLPSDLSQVNV 1085

Qy 1020 -----PPFGIQEERF--NATLKGNSLFDWIKYKTSNHLILVSTAELIENDSVTLPLGQ 1070

Db 1086 SLLLWKTFP-IRAHFSSNLTLRGELK-----SENSSLTSSN----- 1123

Qy 1071 GAFVRSQTEKVEFEFFVNPPLPL--IVGSSVGGILLALLIATAALYKLGFFKRYQKDMSE 1128

Db 1124 --RKREIAIQISKDGLPGRVPLWVILLSAFAGILLMLLILALWKGIFFRPLKMKMEK 1180

RESULT 12

A41131

lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse

N;Alternate names: integrin alpha-4

C;Species: Mus musculus (house mouse)

C;Date: 05-Jun-1992 #sequence\_revision 05-Jun-1992 #text\_change 09-Jul-2004

R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.

J. Cell Biol. 115, 1149-1158, 1991

A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte

A;Reference number: A41131; MUID:92064645; PMID:1840602

A;Accession: A41131

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-1039 <NEU>

A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:G51484; PIDN:CAA37316.1; PID:G51484

C;Superfamily: integrin alpha-4 chain

C;Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 666; DB 2; Length 1039;

Best Local Similarity 22.8%; Pred. No. 5.3e-38;

Matches 283; Conservative 183; Mismatches 405; Indels 368; Gaps 48;

Qy 1 ENLDTENAMTFQ-ENARGFGQSVV-QLOGSR--VVVGAPOEIVAAAN---QRGSLYQCDY 52

Db 41 YNLDPENALLYQPSGTLFGYVVLHSHGSKRWLIUGATASWLSNASVNVNPGAIYRCGI 100

Qy 53 STG---SCFIRLQVP-----VEAVNMSLGLSLAATTPSP-PQLLACG---PTVHQ 95

Db 101 RKNPNQTCQLQSGSPSGPCGKTKCLEERDQNWGLVTLRQPGENGSIIVTCGRWKNIFY 160

Qy 96 TCSNTYVVKLCFLFNSNLRQOPKQKPEALRGCPQSDIAFLIDGSGSIIPHDFRMKE 155

[illegible]

984 TIIITISLLGLIVLLIISCVMMXAGFEKQYKSILOE 1022

Db

RESULT 13  
S06046

integrin alpha-4 chain precursor - human  
N;Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain  
C;Species: Homo sapiens (man)  
C;Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text\_change 09-Jul-2004  
C;Accession: S06046; A39355; D28018  
R;Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.  
EMBO J. 8, 1361-1368, 1989  
A;Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte  
A;Reference number: S06046; MUID:89356603; PMID:2788572  
A;Accession: S06046  
A;Molecule type: mRNA  
A;Residues: 1-1038 <TA>  
A;Cross-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g93945; PIDN:CAA34852.1  
R;Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.  
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991  
A;Title: Characterization of the alpha-4 integrin gene promoter.  
A;Reference number: A39355; MUID:912339513; PMID:2034655  
A;Accession: A39355  
A;Molecule type: DNA  
A;Residues: 1-93 <ROS>  
A;Cross-references: GB:M62841  
R;Takada, Y.; Strominger, J.L.; Hemler, M.E.  
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987  
A;Title: The very late antigen family of heterodimers is part of a superfamily of molecu  
A;Reference number: A94151; MUID:87204112; PMID:3033641  
A;Accession: D28018  
A;Molecule type: protein  
A;Residues: 40-50,'E',52-53 <TA2>  
C;Genetics:  
A;Gene: GDB:ITGA4; CD49D  
A;Cross-references: GDB:128032; OMIM:192975  
A;Map position: 2q31-2q32  
C;Superfamily: integrin alpha-4 chain  
C;Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transmem  
F;1-39/Domain: signal sequence #status predicted <SIG>  
F;40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>

Query Match 10.9%; Score 642; DB 2; Length 1038;  
Best Local Similarity 23.6%; Pred. No. 2.5e-36;  
Matches 293; Conservative 170; Mismatches 408; Indels 368; Gaps 51;

Qy 1 FNLDTENAMTFO-ENARGGQSVV-QLQSR--VVVCAQPIVVAAN---VVGSLYQCDY 52  
Db 40 YNVDTESALLYGGPNTLFGYVVLHSHGANRWLLVGAPTANMLANVASVINPFAIYRCRI 99  
Qy 53 STG---SCEPIRLQVFEAVNMVSLGLSLAATSPPOLLAGPTVHOTCSNTYVKGCLFL 109  
Db 100 GKNPQQTCEQLQSGP-----NCEP---CG---KTCLEERDQWL--- 133  
Qy 110 FGSNLRQOPQKPFPEALRGCPQSDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKT 169  
Db 134 -GVTLSPQGE-----NGSIVTCGHR-----WKN 156  
Qy 170 LFSLMQVSEERFIHTPFKEFQNNPNRSLIKPITQLLRTHATGURKVVRELFNITNGA 229  
Db 157 IFYI----- 160  
Qy 230 RKNAPKILFLLTDGEKFGDPLGYEDVIPELDREGVIR-----YVLGFGDAFSEKSRQ 282  
Db 161 -KNENK---LPTGCYCV-----PDLRTELKRIAPCYQDYVKKGFGENFAS----- 203  
Qy 283 ELNTVASKPRPDHVFQANNFEALKTQVQLREKIFAIEGTQTGSSSSSFEHMSQEGFSAA 342  
Db 204 -----COAGLSSP 211  
Qy 343 ITSNGPLLSTVGSYDWAGGVFLY---TSKEKSTFTINMTVDSMDNDAYLGYA---AAIILR 397  
Db 212 YTKDLIVMGAGSGYWTGSLFVYNTINKYAFLDKQNOVKF---GSYLGYSGVAGHFRS 268



QY 1085 ---FEVFN---PLPLIVG-----SSVGGILLALITAAALYKLGFFKROYKDM 1126  
Db 963 TWFEALHNLPRGVVWIIAISLVLGILFLALLVLLWKGFRFRYKII 1015

RESULT 15  
T31437  
integrin alpha chain su2 - sea urchin (Lytechinus variegatus)  
C;Species: Lytechinus variegatus (variegated urchin)  
C;Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T31437  
R;Hertzler, P.L.; McClay, D.R.  
submitted to the EMBL Data Library, May 1998  
A;Description: Alpha su2, a sea urchin integrin which binds laminin.  
A;Reference number: Z21035  
A;Accession: T31437  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1041 <HER>  
A;Cross-references: UNIPROT:O76378; EMBL:AF067658; NID:g3220241; PIDN:AAC2  
A;Experimental source: developmental stage embryo  
C;Function:  
A;Description: binds laminin  
C;Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;  
Best Local Similarity 24.8%; Pred. No. 2.2e-34;  
Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

QY 336 QEGFSAITNGP--LLSTVGSYDAGGVFLYTSKESKSTFNTRVDS----DMNDAYLG 389  
Db 181 QAGFGIIFSDNSALVAGFSGYVLOQIYVQSLLNRSV-VQATQESNTGYSFDSYRG 239

QY 390 YAAAI--ILNRVQSLVLAGAPRYQHI--GLVAMFRONTGMWESNANVKGTGIGAYFGASLC 446  
Db 240 YSLALGDFNGDGVQDYVVGTPRAESLMGLVAIFDQNLQFN---QVMGTQIVAYFGYSVT 296

QY 447 SVDVDNGSTDLVLIGAPHYEQTRGGQSVCEPLRGQRARWOCDAVLYGEQ----- 498  
Db 297 VVDI--NNDTVDDLVLGAPWYMDGPAIQ-----RWAGAVVYLQNPVGPGA 343

QY 499 -----QPWGRFGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLFHGTSGS 546  
Db 344 SNRLSLSSTLIGQIRSRFGLSIAGSDSNODQFNDVAIGAPYEGDDAGAVYIYG-SAN 402

QY 547 GTSFHSQRIAGSKLS-PRQXFGQSLSGODITMDGLVDLTVGAQ--GHVLLRSQPVL 603  
Db 403 GLKSTPAQVLTPTLGHSGITTFGSLQGQDMKNKYPDLLVGAESANTAVLIRTPV 462

QY 604 RVKAIMFNPREVARNVFECNDQVWKGAGEVRVCLHVQKSTRDRIREGOISV----- 658  
Db 463 SLDATLNTFP-----IGINLENKTYE-LADGTWVTSFIAMT 497

QY 659 -----VTYDIALDSG-RPHSRVFNETHKSNTRRQTVLGL-TQTCETKLQ 702  
Db 498 CFTYTGNYLPDHIDISYTVTVDSGIITANRRAMPVNDMSEITKTRRLAVSTQFCDFL 557

QY 703 LPNCIEDPVSPIVLRNLSLVGTPLSAFCN-----LRPVLAEADAORLFTALFPPEK 753  
Db 558 VGNSTIEDKLTPKVTLOYDL-----NNDERSRIQPHILPIIDMATWMTSTQTKVSIQ 609

QY 754 NCGNINICODDLSITFSFMSLDCVLVGGPREFNVTVRNDCGEDSVYRTOVTFEFLDLSY 813  
Db 610 NCVN-NICIPDLQVTVT-FNLNIVIGQVQLTLDVSLNRRGEDAFQSSLSVYPLGLQF 667

QY 814 RNVSTLQNRORSORWRLACESASSTEVSGALKSTGCSINHPIFPPN-----SEVTFNIT 867  
Db 668 VRL-----ERKANMDFSVTCSDS-----LRITCDTGNPMVGNKILFGLTSTFQVS 717

QY 868 FDVDS-----KASLGNKULLXANVTSENMPKNTKTEFQELPVKYAVVMVTVTSHGVSTK 922  
Db 718 GDKDSIEFYFKAESNS--EDPNTLENNELNMVTPVTDCTLKLLSASYPEIVMYSTQED 775

Search completed: November 9, 2004, 12:19:22  
Job time : 34.25 secs

QY 923 YL-----NETASENTSRVMOHOYQV-----SNLQGRSLPIS-----LVFLVPV 960  
Db 776 YVVPFPFPAKNAEADIGMEVN-HLYEVRNTGSSNAGEVSLNIQWPQKNEDEYLFYLLGI 834

QY 961 RLNGTVIWRDPO-----VTFSENLSSTCHTKERLPSPSHSDFLAELRKAPVV 1005  
Db 835 MTEGVTCLTQKCANPEGVKLEPSTKAKJNSNTTQVSGRKRREPEVAEALAQTDN--VI 892

QY 1006 NCSTAVCQRIQCDIFFFGIOEEFNAT-----LKGNLSEFDWYIKTSHNHLIVSTAELF 1059  
Db 893 YCASDSCVLINCTI-----DEINASKKVRIILGRF-----W-----ERTF 928

QY 1060 NDSVFTLLPGQGFVRQSTETKVE--PFEVPNP-----LP----- 1092  
Db 929 QKAVSELTPTVVOATIASASAANKTIYNIPLPRDFSDSTKASTLVTTTELVPVPTPIAW 988

QY 1093 -LIVGSSVGGILLALITAAALYKLGFFKR-----QYKDMMS--EGGPP 1132  
Db 989 WIIVSVLGGIILILLIILGLMKGFFERKKPGKEEYAPVASADKGGPP 1038



GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds  
(without alignments)  
4783.919 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARFGQ.....FKROYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	5823.5	99.1	1152	1 ITAM HUMAN	P11215 homo sapien
2	4447	75.6	1153	1 ITAM MOUSE	P05555 mus musculus
3	4358	74.1	1151	2 Q9J130	Q9J130 rattus norv
4	3786.5	64.4	920	2 Q28984	Q28984 sus scrofa
5	3456	58.8	1163	1 ITAX HUMAN	P20702 homo sapien
6	3388	57.6	1162	1 ITAD HUMAN	Q13349 homo sapien
7	3306.5	56.2	1188	2 Q6KAS4	Q6KAS4 mus musculus
8	3306.5	56.2	1188	2 Q6KAS4	Q6KAS4 mus musculus
9	3297.5	56.1	1169	1 ITAX MOUSE	Q9GKH4 mus musculus
10	3225.5	54.9	1161	1 ITAD RAT	Q9GKH4 mus musculus
11	1561	26.6	1165	1 ITAL BOVIN	P61625 bos taurus
12	1539	26.2	1166	2 Q6TVB8	Q6TVB8 bos taurus
13	1539	26.2	1166	2 AAQ90015	AAQ90015 bos tauru
14	1536.5	26.1	1170	1 ITAL HUMAN	P20701 homo sapien
15	1516.5	25.8	1163	1 ITAL MOUSE	P24063 mus musculus
16	1512.5	25.7	1161	2 Q9WTV4	Q9WTV4 mus musculus
17	1502	25.5	1160	2 Q9R200	Q9R200 mus musculus
18	1382	23.5	1196	2 Q9RTF1	Q9RTF1 cyprinus ca
19	1341.5	22.8	1086	2 Q9GHB1	Q9GHB1 homo sapien
20	1329.5	22.6	1187	2 Q9RTF0	Q9RTF0 cyprinus ca
21	1264	21.5	927	2 Q8HZV0	Q8HZV0 bos taurus
22	1142.5	19.4	1167	2 Q8B340	Q8B340 rattus norv
23	1136.5	19.3	1167	1 ITAE MOUSE	Q8B340 rattus norv
24	1128	19.2	1179	1 ITAE HUMAN	P38570 homo sapien
25	1097	18.7	1167	2 Q8B341	Q8B341 rattus norv
26	1085.5	18.5	1151	1 ITAL HUMAN	P56199 homo sapien
27	1079.5	18.4	1189	1 ITAH HUMAN	Q9UKX5 homo sapien
28	1071	18.2	1170	1 ITA2 BOVIN	P51710 bos taurus
29	1057	18.0	1178	1 ITA2 MOUSE	Q62469 mus musculus
30	1057	18.0	1178	2 Q6P1C7	Q6P1C7 mus musculus
31	1057	18.0	1178	2 AAH65139	AAH65139 mus muscu

## ALIGNMENTS

## RESULT 1

ID	ITAM HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein Mac-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor M01)			
DE	(Neutrophil adherence receptor).			
GN	Name=ITGAM; Synonyms=CR3A, CD11B;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Artiodactyla; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			

Proc. Natl. Acad. Sci. U.S.A. 86:257-261 (1989).  
[6]  
RN MEDLINE=87076671; PubMed=353202; Todd R.F. III, Arnaut M.A.;  
RP SEQUENCE OF 1-9 FROM N.A.  
RX TISSUE=BLOOD;  
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;  
RT "The promoter of the CD11b gene directs myeloid-specific and  
RT developmentally regulated expression.";  
Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529 (1991).  
[7]  
RN MEDLINE=92144986; PubMed=1345576;  
RP SEQUENCE OF 1-9 FROM N.A.  
RX TISSUE=BLOOD;  
RA Pahl H.L., Rosmarin A.G., Tenen D.G.;  
RT "Characterization of the myeloid-specific CD11b promoter.";  
Blood 79:865-870 (1992).  
[8]  
RN MEDLINE=87076671; PubMed=353202;  
RP SEQUENCE OF 17-31.  
RX MEDLINE=87076671; PubMed=353202; Todd R.F. III, Arnaut M.A.;  
RA Pierce M.W., Remold-O'Donnell E.,  
RT "N-terminal sequence of human leukocyte glycoprotein Mol: conservation  
RT across species and homology to platelet Iib/IIia.";  
Biochim. Biophys. Acta 874:368-371 (1986).  
[9]  
RN X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.  
RP MEDLINE=95171458; PubMed=7867070;  
RX Lee J.O., Rieu P., Arnaut M.A., Liddington R.;  
RA "Crystal structure of the A domain from the alpha subunit of integrin  
RT CR3 (CD11b/CD18).";  
Cell 80:631-638 (1995).  
[10]  
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 148-334.  
RP MEDLINE=96363671; PubMed=8747460;  
RX Lee J.O., Bankston L.A., Arnaut M.A., Liddington R.C.;  
RA "Two conformations of the integrin A-domain (I-domain): a pathway for  
RT activation?";  
Structure 3:1333-1340 (1995).  
[11]  
RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 148-337.  
RP MEDLINE=98362595; PubMed=9687375;  
RX Baldwin E.T., Sarver R.W., Bryant G.L. Jr., Curry K.A.,  
RA Fairbanks M.B., Finzel B.C., Garlick R.L., Heinrikson R.L.,  
RA Horton N.C., Kelley L.L., Mildner A.M., Moon J.B., Mott J.E.,  
RA Mutchler V.T., Tomich C.S., Watenpaugh K.D., Wiley V.H.;  
RT "Cation binding to the integrin CD11b I domain and activation model  
RT assessment.";  
Structure 6:923-935 (1998).  
[12]  
RN 3D-STRUCTURE MODELING OF 17-616.  
RP MEDLINE=98226734; PubMed=9560195;  
RX Oxvig C., Springer T.A.;  
RA "Experimental support for a beta-propeller domain in integrin alpha-  
RT subunits and a calcium binding site on its lower surface.";  
Proc. Natl. Acad. Sci. U.S.A. 95:4870-4875 (1998).  
CC -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various  
CC adhesive interactions of monocytes, macrophages and granulocytes  
CC as well as in mediating the uptake of complement-coated particles.  
CC It is identical with CR-3, the receptor for the iC3b fragment of  
CC the third complement component. It probably recognizes the R-G-D  
CC peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for  
CC fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides  
CC of fibrinogen gamma chain.  
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-M  
CC associates with beta-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and  
CC granulocytes.  
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
CC with I-domains do not undergo protease cleavage.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11b entry;  
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11b.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J03925; AAA59544.1; -  
CC EMBL; M18044; AAA59491.1; -  
CC EMBL; J04145; AAA59903.1; -  
CC EMBL; S52227; AAB24821.1; -  
CC EMBL; S52152; AAB24821.1; JOINED.  
CC EMBL; S52153; AAB24821.1; JOINED.  
CC EMBL; S52154; AAB24821.1; JOINED.  
CC EMBL; S52155; AAB24821.1; JOINED.  
CC EMBL; S52157; AAB24821.1; JOINED.  
CC EMBL; S52159; AAB24821.1; JOINED.  
CC EMBL; S52161; AAB24821.1; JOINED.  
CC EMBL; S52164; AAB24821.1; JOINED.  
CC EMBL; S52165; AAB24821.1; JOINED.  
CC EMBL; S52167; AAB24821.1; JOINED.  
CC EMBL; S52169; AAB24821.1; JOINED.  
CC EMBL; S52170; AAB24821.1; JOINED.  
CC EMBL; S52173; AAB24821.1; JOINED.  
CC EMBL; S52174; AAB24821.1; JOINED.  
CC EMBL; S52180; AAB24821.1; JOINED.  
CC EMBL; S52181; AAB24821.1; JOINED.  
CC EMBL; S52184; AAB24821.1; JOINED.  
CC EMBL; S52189; AAB24821.1; JOINED.  
CC EMBL; S52191; AAB24821.1; JOINED.  
CC EMBL; S52192; AAB24821.1; JOINED.  
CC EMBL; S52203; AAB24821.1; JOINED.  
CC EMBL; S52212; AAB24821.1; JOINED.  
CC EMBL; S52213; AAB24821.1; JOINED.  
CC EMBL; S52216; AAB24821.1; JOINED.  
CC EMBL; S52219; AAB24821.1; JOINED.  
CC EMBL; S52220; AAB24821.1; JOINED.  
CC EMBL; S52221; AAB24821.1; JOINED.  
CC EMBL; S52222; AAB24821.1; JOINED.  
CC EMBL; S52226; AAB24821.1; JOINED.  
CC EMBL; M76724; AAA58410.1; -  
CC EMBL; M84477; AAA51960.1; -  
CC PIR; A31108; RWHU1B.  
CC PDB; 1A8X; Model; @=17-1152.  
CC PDB; 1BHO; X-ray; 1/2=-.  
CC PDB; 1BHQ; X-ray; 1/2=-.  
CC PDB; 1IDN; X-ray; 1/2=-.  
CC PDB; 1IDO; X-ray; @=140-331.  
CC PDB; 1JLM; X-ray; @=143-334.  
CC PDB; 1MIU; X-ray; A=137-331.  
CC PDB; 1MF7; X-ray; A=144-337.  
CC PDB; 1N92; X-ray; A=140-335.  
CC PDB; 1NA5; X-ray; A=144-345.  
CC Genew; HGNC:6149; ITGAM.  
CC MIM; 120980; -  
CC GO; GO:0008305; C:integrin complex; TAS.  
CC GO; GO:0007155; P:cell adhesion; TAS.  
CC GO; GO:0007043; Integrin\_alpha.  
CC InterPro; IPR000413; Integrin\_alpha.  
CC InterPro; IPR002035; VWFA.  
CC Pfam; PF01839; FG-GAP; 3.  
CC Pfam; PF00357; Integrin\_alpha; 1.  
CC Pfam; PF00092; VWFA; 1.  
CC PRINTS; PR01195; INTEGRINA.  
CC SMART; SM00191; Int\_alpha; 5.  
CC SMART; SM00327; VWFA; 1.  
CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
CC PROSITE; PS0234; VWFA; 1.  
CC 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;  
CC Glycoprotein; Integrin; Magnesium; Receptor; Repeat; Signal;  
CC Transmembrane.



[illegible]

RT complement receptor type 3 and cellular adhesion molecule Mac-1." ;  
Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).  
[4]  
RN MEDLINE=95188276; PubMed=3887182;  
RX Springer T.A., Teplov D.B., Dreyer W.J.;  
RA "Sequence homology of the IFA-1 and Mac-1 leukocyte adhesion  
RT glycoproteins and unexpected relation to leukocyte interferon." ;  
Nature 314:540-542(1985).  
RL  
CC -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various  
CC adhesive interactions of monocytes, macrophages and granulocytes  
CC as well as in mediating the uptake of complement-coated particles.  
CC It is identical with CR-3, the receptor for the IC3b fragment of  
CC the third complement component. It probably recognizes the R-G-D  
CC peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for  
CC fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides  
CC of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in  
CC mast cell development and in immune complex-mediated  
CC glomerulonephritis. Mice expressing a null mutation of the alpha-M  
CC subunit gene demonstrate increase in neutrophil accumulation, in  
CC response to a impaired degradation and phagocytosis, events that  
CC apparently accelerate apoptosis in neutrophils. These mice develop  
CC obesity. Heterodimer of an alpha and a beta chain. Alpha-M  
CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M  
CC associates with beta-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=P05555-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=P05555-2; Sequence=VSP\_010473;  
CC Note=No experimental confirmation available;  
CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and  
CC granulocytes  
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
CC with I-domains do not undergo protease cleavage. family.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; X07640; CAA30479.1; -.  
CC EMBL; AK039444; BAC30350.1; -.  
CC EMBL; M14293; AAA39484.1; -.  
CC FIC; S00551; S00551.  
CC HSP; P11215; LBHQ.  
CC MGD; MGI:96607; Icam.  
CC GO; GO:0009897; C:external side of plasma membrane; IDA.  
CC GO; GO:0007155; P:cell adhesion; IMP.  
CC GO; GO:0045123; P:cellular extravasation; IMP.  
CC GO; GO:0030593; P:neutrophil chemotaxis; IMP.  
CC InterPro; IPR000413; Integrin\_alpha.  
CC InterPro; IPR002035; VWF\_A.  
CC Pfam; PF01839; FG-GAP; 3.  
CC Pfam; PF00357; Integrin\_alpha; 1.  
CC Pfam; PF00092; VWFA; 1.  
CC PRINTS; PR01185; INTEGRINA.  
CC PRINTS; PR00453; VWFADOMAIN.  
CC SMART; SM00191; Int\_alpha; 5.  
CC SMART; SM00327; VWFA; 1.  
CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
CC PROSITE; PS0234; VWFA; 1.  
CC Alternative splicing; Calcium; Cell adhesion;  
CC Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;  
CC Signal; Transmembrane.

FT	SIGNAL	1	16
FT	CHAIN	17	1153
FT	DOMAIN	17	1105
FT	TRANSMEM	1106	1129
FT	DOMAIN	1130	1153
FT	REPEAT	31	84
FT	REPEAT	85	163
FT	DOMAIN	164	350
FT	REPEAT	337	400
FT	REPEAT	401	452
FT	REPEAT	454	515
FT	REPEAT	517	575
FT	REPEAT	580	632
FT	CA_BIND	465	473
FT	CA_BIND	529	537
FT	CA_BIND	592	600
FT	SITE	1132	1136
FT	DISULFID	66	73
FT	DISULFID	105	123
FT	DISULFID	654	711
FT	DISULFID	770	776
FT	DISULFID	999	1023
FT	DISULFID	1028	1033
FT	CARBOHYD	58	58
FT	CARBOHYD	86	86
FT	CARBOHYD	391	391
FT	CARBOHYD	696	696
FT	CARBOHYD	734	734
FT	CARBOHYD	772	772
FT	CARBOHYD	801	801
FT	CARBOHYD	881	881
FT	CARBOHYD	907	907
FT	CARBOHYD	941	941
FT	CARBOHYD	980	980
FT	CARBOHYD	994	994
FT	CARBOHYD	1022	1022
FT	CARBOHYD	1045	1045
FT	CARBOHYD	1051	1051
FT	CARBOHYD	1076	1076
FT	VARSPLIC	453	569
FT	CONFLICT	37	37
FT	CONFLICT	683	683
FT	SEQUENCE	1153 AA;	127480 MW; 178DS988AECB0343 CRC64;
FT	SEQUENCE	1153 AA;	127480 MW; 178DS988AECB0343 CRC64;

Query Match 75.6%; Score 4447; DB 1; Length 1153;  
Best Local Similarity 73.5%; Pred. No. 9,7e-287;  
Matches 836; Conservative 147; Mismatches 153; Indels 2; Gaps 2;

QY	1	FNLDTENAMTFOENARGGQSVVQLGSRVVVGAPQEIIVANQGRSLYOCDYSTGSCPEI	60
DB	17	FNLDTHEPMTFOENAKGFGQNVVQLGGTSVVVAAPQAKAVNQTGALYQCDYSTSRCHP	76
QY	61	RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHTCSENTYVKGICFLFGSNLRQPOK	120
DB	77	PLQVPPPEAVNMSLGLSLAVSTVPQOLLACGPTVHTCSENTYVKGICFLFGSNLRQPOK	136
QY	121	FPEARLGCQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSEEF	180
DB	137	FPEARLGCQEDSDIVFLIDGSGSINIDFQKKKEFVSVMEQFKSKTLFSLMOYSDEF	196
QY	181	RIHFTPFKEPQNNPNRSLIKPIQLLGRTHTTATGLKVVVRELFNITNGARKNAFKILFL	240
DB	197	RIHFTFNDKPNRSPRSHVSPKQLNGRTKTASGIRKVVRELPHKNGARENAKTLVVI	256
QY	241	TGGEKFGDPLGYEDVITPELDRECVIRYVLGFGDARSEKSRQELNVTASPPRDHVFQAN	300
DB	257	TGGEKFGDPLGYEDVITPELDRECVIRYVLGFGDARSEKSRQELNVTASPPRDHVFQAN	316
QY	301	NFEALKTQVNLREKIFAEJGTGTGSSSSFEHMSQEGFSAATITSGNPLLSVGVSDWAG	360
DB	317	NFEALNTIQNLQEKIFAEJGTGTGSSSSFEHMSQEGFSAATITSGNPLLSVGVSDWAG	376

```

QY 361 GVFLYTSKEKSTFNNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGASAVILNRVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKTGICAGYFASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSCPL 480
Db 437 ENFGTWESHTSISGQISYFASLCSVDMDADGNTNLILGAPHYEQTRGGQVSCPL 496
QY 481 PRGORARQCDVAVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVILF 540
Db 497 PRG-RARWQCEALLHGDGHPWGRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVILF 555
QY 541 HGTSGSGLSPHSQRIAGSKLSPRLQYFGQSLSGQDGLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 YGASIASLSASHSHRIIAGHSPGLQYFGQSLSGQDGLTMDGLVDLTVGAQGHVLLRSQ 615
QY 601 PVLRVKAIMENPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDRREGIOQSVVT 660
Db 616 PVLREATMEFSPKKVARSVFACQEVLLKNKDAGEVRVCLVRKNTKDLREGDIQSTVT 675
QY 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLGTOTCETLKLQLENCLIEDPSPVILRLNF 720
Db 676 YDLALDPVRSIRAFDETKNTRRTQVFGLMQKCETLKLILPDCVDDSVSPILRLNY 735
QY 721 SLVGTPLSAFNGURPVLAEDAQRALTALPFPKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 TLVGEPLRSFGLRPVLAMDARFFETAMFPFKNCNDNICQDDLSITFSFMSLDCLVVG 795
QY 781 GPREENVTVRNDGDSYTOTVTFPPDLSTRKUSTLQNRQSRWRL-ACESASSTE 839
Db 796 GPQDFNMSVTRNDGDSYTOTVTFPPDLSTRKUSTLQNRQSRWRL-ACESASSTE 855
QY 840 VSGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMWRINKTE 899
Db 856 GHGALKSTTWINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMWRINKTE 915
QY 900 FQLELPVKAVYVNTSHGVSTKYLNTFASENTSRVMOHQYQVSNLQORSLPISLVFLVP 959
Db 916 FQLELPVKAVYVNTSHGVSTKYLNTFASENTSRVMOHQYQVSNLQORSLPISLVFLVP 975
QY 960 VRLNQIVDRPQVTSSENLSSTCHTKERLPSHDSFLAELRKAPVNCSTIACQRIQCDI 1019
Db 976 VQLNNVTVMHPQVIFESQNLSSACHTEQSPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035
QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNLLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
Db 1036 PSFNTQEIENVTLKNLSFDWYIKTSHNLLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1095
QY 1080 TKVEPEVNPPLIVGSSVGGLLLALITALYKLGFFKRYQKMMWSEGGPPGAPQ 1137
Db 1096 TKVEPEVNPPLIVGSSVGGLLLALITALYKLGFFKRYQKMMWSEGGPPGAPQ 1153

RESULT 3
Q9J130
ID Q9J130 PRELIMINARY; PRT; 1151 AA.
AC 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DE 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zerria K. Jr.
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AF268593; AAF81280.1; -.
DR HSP; F11215; I8HQ.

```

```

DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 2.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS02034; VWFA; 1.
DR Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.1%; Score 4358; DB 2; Length 1151;
Best Local Similarity 72.3%; Pred. No. 8.1e-281;
Matches 822; Conservative 151; Mismatches 162; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARGGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQCPQK 120
Db 77 FLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSENTYVKGCLFLFGSNLRQCPQK 136
QY 121 FPEALGCCQEDSDIAFLIDGSSIIIPHRMKELVSTIMBOLKXKTLFSLMOYSEEF 180
Db 137 FPEALGCCQEDSDIAFLIDGSSIIIPHRMKELVSTIMBOLKXKTLFSLMOYSEEF 196
QY 181 RIHFTKEKFNPNRSLIKPITOLLGRTHATGLRKYVRELNIINGARKNAKFLFL 240
Db 197 RIHFTKEKFNPNRSLIKPITOLLGRTHATGLRKYVRELNIINGARKNAKFLFL 256
QY 241 TDGEKFGDPLGYEDVPIELDREGVIRYVLGDFGDAFSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDGEKFGDPLGYEDVPIELDREGVIRYVLGDFGDAFSEKSRQELNTVASKPPRDHVFQAN 316
QY 301 NFEALKTQVONREKIFALEGTQSGSSSEFHEMSQEGFSAATISNGPLISTVGSYDWAG 360
Db 317 NFEALKTQVONREKIFALEGTQSGSSSEFHEMSQEGFSAATISNGPLISTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFNNTRVDSMDNDAYLGAAAIILNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGASAVILNRVQSLVLCAPRYQHIGLVAMFR 436
QY 421 QNTGWESNANVKTGICAGYFASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSCPL 480
Db 437 ENFGTWESHTSISGQISYFASLCSVDMDADGNTNLILGAPHYEQTRGGQVSCPL 496
QY 481 PRGORARQCDVAVLYGEGQCPWGRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVILF 540
Db 497 PRG-RARWQCEALLHGDGHPWGRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVILF 555
QY 541 HGTSGSGLSPHSQRIAGSKLSPRLQYFGQSLSGQDGLTMDGLVDLTVGAQGHVLLRSQ 600
Db 556 YGASIASLSASHSHRIIAGHSPGLQYFGQSLSGQDGLTMDGLVDLTVGAQGHVLLRSQ 615
QY 601 PVLRVKAIMENPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDRREGIOQSVVT 660
Db 616 PVLREATMEFSPKKVARSVFACQEVLLKNKDAGEVRVCLVRKNTKDLREGDIQSTVT 675
QY 661 YDLALDSGRPHSRVAFNETKNSRTRQVGLGTOTCETLKLQLENCLIEDPSPVILRLNF 720
Db 676 YDLALDPVRSIRAFDETKNTRRTQVFGLMQKCETLKLILPDCVDDSVSPILRLNY 735
QY 721 SLVGTPLSAFNGURPVLAEDAQRALTALPFPKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 TLVGEPLRSFGLRPVLAMDARFFETAMFPFKNCNDNICQDDLSITFSFMSLDCLVVG 795

```





FT TURN 160 161  
 FT HELIX 164 178  
 FT TURN 179 180  
 FT TURN 183 185  
 FT STRAND 186 191  
 FT STRAND 197 201  
 FT HELIX 203 208  
 FT HELIX 212 216  
 FT TURN 217 218  
 FT STRAND 226 226  
 FT HELIX 228 236  
 FT TURN 237 240  
 FT HELIX 242 244  
 FT TURN 245 245  
 FT TURN 248 249  
 FT STRAND 251 258  
 FT STRAND 263 263  
 FT HELIX 269 278  
 FT TURN 279 280  
 FT STRAND 282 288  
 FT HELIX 290 293  
 FT TURN 296 297  
 FT HELIX 298 304  
 FT HELIX 310 312  
 FT STRAND 313 316  
 FT HELIX 319 325  
 FT HELIX 326 334  
 FT TURN 335 335  
 SQ SEQUENCE 1163 AA; 127886 MW; 83658A13B5C5D8F CRC64;

Query Match 58.8%; Score 3456; DB 1; Length 1163;  
 Best Local Similarity 60.6%; Pred. No. 9,3e-221;  
 Matches 684; Conservative 143; Mismatches 296; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGGRRVVVGAPQEIIVAAANQKSLYQCDYSTGSCBPI 60  
 DB 20 FNLDTEELTAFRVDSAGDSVVQYANSVVVGAPQKITAANQTCGLYQCGYSTGACBPI 79  
 QY 61 RLQVPEAVNMSLGLSLAATPPQLLACGPTVHTQCSNTYVKGCLFLFGSNLQQOQK 120  
 DB 80 GLQVPEAVNMSLGLSLASTTSPQLLACGPTVHHECGRNMYLTGLCFLGPT--QLTOR 137  
 QY 121 FPEARLGCQEDDIAFLDGGSSIPHPDFRMKELVSTIMEQLKSKTLFSLMOYSEEF 180  
 DB 138 LPVSRQECPRQODIVFLDGGSSISRNFATMNFVRAVISQFORPSTQFSLMQFSNKF 197  
 QY 181 RIHFTFKPQNNPNRSLKIPITQLLGRTHRTATGLRKKVYRELFTNNGARKNAKILFLL 240  
 DB 198 QTHFTFEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257  
 QY 241 TDGKFGDPLGYEDVIPLEDRGVIRYVILGFGDAFRSEKSRQELNTVASKPRDHFVQAN 300  
 DB 258 TDGKFGDPLGYEDVIPLEDRGVIRYVILGFGDAFRSEKSRQELNTVASKPRDHFVQAN 317  
 QY 301 NFEALKTVOQLREKIFATEGTQTSSSSFEHEMSQESAAITNSGPLLSTVGSYDWAG 360  
 DB 318 DFDALNDIQNKEKIFALEGHTTSSSFEHEMSQESAAITNSGPLLSTVGSYDWAG 377  
 QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILNLRVQSLVILGAPRYQHIGLVAMPR 420  
 DB 378 GAFLYPPNMSPTFINMSQENVDNRDSYLGYSTELALWKGVSILVILGAPRYQHTGKAVIET 437  
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDSNGSTDVLVILGAPRYQHTGKAVIET 480  
 DB 438 QVSRQWRMAEAVTGTQISYFGASLCSDVDSNGSTDVLVILGAPRYQHTGKAVIET 497  
 QY 481 PRGORARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
 DB 498 PRGWR-RWWCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVIGAPGEEDNRGAVYLF 556  
 QY 541 HGTSGGSGPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQHVLIRSQ 600  
 DB 557 HGVILGPSISPSHSQRIAGSKLSSRLQYFGQSLSGGQDLTMDGLVDLTVGAQHVLIRSQ 616

QY 601 PVLVRKAIMBENPREVARNVFECNDQVYVKGKEAGEVVRVCLHVQKSTDRLEREGIOISVVT 660  
 DB 617 PVLVWVGSMQFIPAEIIPRSAPFECREQVVSQTLVQSNICLYIDKRSKNLGGSRDLQSSVT 676  
 QY 661 YDLALDSGRPHSRAVFNETKNSRTRQVTLGTCTCETLKLQIPNCIEDPVSPIVLRNF 720  
 DB 677 LDIALDPGRLSPRATFOETKNSRSLRVVLGLKAHCENFNLLLESCVEDSVTPITLRNF 736  
 QY 721 SLVCTPLSARGNLRPLVLAEDAQRLLFTALFPFCKNCGNDNICQDDLSITFFSFMSDCLVVG 780  
 DB 737 TLVGKPLLAFLRNLRPLAADAQRYFTASLPFCKNCGADHICQDNLGTSFSPGLKSLVVG 796  
 QY 781 GPREFNVTVTRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSRWRLACESASSTEV 840  
 DB 797 SNLENAEAVWVWVNDGEDSYGTTITFSHPAGLSYVVAEGQKQGLRSLHLTCDSPAVG-- 854  
 QY 841 SGALKTSCTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPKNTKEF 900  
 DB 855 SQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVLGDRLLLTANVSSENTPRTSKTTF 914  
 QY 901 QLELPVYAVVWVVTSHGVSTKYNFTAS-ENTSRVMOHQYQVSNLQORSLPISLVFLVP 959  
 DB 915 QLELPVYAVVWVVTSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQORDLPVSINFVWP 974  
 QY 960 VRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVWVNGSIAVCQRIQCDI 1019  
 DB 975 VELNQEAVMVDVEVSHPCNSLRCSEKIAAPPASDFLAHIQKNPVLDCSIAGCLFRCDV 1034  
 QY 1020 PFFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAETLFDNSVFTLAPGGQAFVRSOTE 1079  
 DB 1035 PSFSVQEBLDTLKNLSFGWVRQLQKKVSVSWAEITFDTSVYSQPLPGQAFVRAQT 1094  
 QY 1080 TKVPEFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMESE 1128  
 DB 1095 TVLEKIKVHNPTPLIVGVSSIGLLLLALITAVLYKVGFFKQYKEMMEE 1143

RESULT 6  
 ITAD HUMAN  
 ID Q13349; Q15575; Q15576; PRT; 1162 AA.  
 AC 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).  
 GN Name=ITGAD;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]\_TaxID=9606;  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Splice;  
 RX MEDLINE=96111956; PubMed=8777714;  
 RA Van der Vieren M., Le Trong H., Wood C.L., Moore P.F., St John T.,  
 RA Staunton D.E., Gallatin W.M.;  
 RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-  
 RT 3";  
 RL Immunity 3:683-690(1995).  
 RN [2]  
 RP SEQUENCE OF 1-235 FROM N.A.  
 RX MEDLINE=20187620; PubMed=10722744;  
 RA Noti J.D., Johnson A.K., Dillon J.D.;  
 RT "Structural and functional characterization of the leukocyte integrin  
 RT gene CD11d. Essential role of Sps and Sp3";  
 RL J. Biol. Chem. 275:8959-8969(2000).  
 RN [3]  
 RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.  
 RX MEDLINE=96257236; PubMed=8666289;  
 RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;  
 RT "Cloning and chromosomal localization of a novel gene-encoding a human  
 RT beta 2-integrin alpha subunit.";

RL Gene 171:291-294 (1996).  
 RN [4]  
 RP INTERACTION WITH VCAM1.  
 RX MEDLINE=99059842; PubMed=9841932;  
 RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,  
 RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
 RT "alpha2beta2 integrin is expressed on human eosinophils and functions  
 as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-  
 1).";  
 RL J. Exp. Med. 188:2187-2191 (1998).  
 RN [5]  
 RP INTERACTION WITH VCAM1.  
 RX MEDLINE=99370002; PubMed=1043935;  
 RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
 RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
 RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
 binding interface between I domain and VCAM-1";  
 RL J. Immunol. 163:1984-1990 (1999).  
 CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and  
 CC VCAM1. May play a role in the atherosclerotic process such as  
 CC clearing lipoproteins from plaques and in phagocytosis of blood-  
 CC borne pathogens, particulate matter, and senescent erythrocytes  
 CC from the blood.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D  
 CC associates with beta-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell  
 CC lines and subsets of peripheral blood leukocytes and strongly on  
 CC tissue-specialized cells, including macrophages foam cells within  
 CC atherosclerotic plaques, and on splenic red pulp macrophages.  
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U37028; AAB38547.1; -  
 DR EMBL; U40274; AAB60634.1; -  
 DR EMBL; U40275; AAB60635.1; -  
 DR EMBL; U40276; AAB60636.1; -  
 DR EMBL; U40277; AAB60637.1; -  
 DR EMBL; U40279; AAB60638.1; -  
 DR EMBL; U40278; AAB60638.1; JOINED.  
 DR EMBL; AF187881; AAF62875.1; -  
 DR HSSP; P11215; IBHQ.  
 DR Genew; HGNC:6146; ITGAD.  
 DR MIT; 602453; -  
 DR GO; GO:0008305; C:integrin complex; TAS.  
 DR GO; GO:0016337; P:cell-cell adhesion; NAS.  
 DR GO; GO:0007160; P:cell-matrix adhesion; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWA; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00327; VWA; 1.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS0234; VWFA; 1.  
 KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 17 potential.

FT CHAIN 18 1162 Integrin alpha-D.  
 FT DOMAIN 18 1100 Extracellular (Potential).  
 FT TRANSMEM 1101 1121 Potential.  
 FT DOMAIN 1122 1162 Cytoplasmic (Potential).  
 FT REPEAT 32 85 FG-GAP 1.  
 FT REPEAT 86 75 FG-GAP 2.  
 FT DOMAIN 150 332 VWFA.  
 FT REPEAT 350 400 FG-GAP 3.  
 FT REPEAT 401 452 FG-GAP 4.  
 FT REPEAT 454 516 FG-GAP 5.  
 FT REPEAT 518 576 FG-GAP 6.  
 FT REPEAT 581 633 FG-GAP 7.  
 FT CA\_BIND 465 473 Potential.  
 FT CA\_BIND 530 538 Potential.  
 FT CA\_BIND 593 601 Potential.  
 FT SITE 1127 1131 GFEKR motif.  
 FT DISULFID 67 74 By similarity.  
 FT DISULFID 106 124 By similarity.  
 FT DISULFID 655 710 By similarity.  
 FT DISULFID 769 775 By similarity.  
 FT DISULFID 846 861 By similarity.  
 FT DISULFID 994 1018 By similarity.  
 FT DISULFID 1023 1028 By similarity.  
 FT CARBOHYD 59 59 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 99 99 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 391 391 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 691 691 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 733 733 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 873 873 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 957 957 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 1046 1046 N-linked (GlcNAc...) (Potential).  
 FT CONFLICT 500 500 Missing (in Ref. 2).  
 FT CONFLICT 515 518 GHPW -> ATP (in Ref. 2).  
 FT CONFLICT 825 825 L -> V (in Ref. 2).  
 FT CONFLICT 984 984 V -> A (in Ref. 2).  
 SQ SEQUENCE 1162 AA; 126885 MW; F296AIA35455D77D CRC64;  
 Query Match 57.6%; Score 3388; DB 1; Length 1162;  
 Best Local Similarity 58.9%; Pred. No. 3.1e-216;  
 Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;  
 QY 1 FNLDTENAMTFOENARFGSGVVLQSGRVVGVGAPQBIIVANORGSILYQCDYSGSCPI 60  
 DB 18 FNLDVEEPTIFQEDAGFGSGVVLQSGRVVGVGAPQBIIVANORGSILYQCDYSGSCPI 77  
 QY 61 RLQVPVAVNNSLGLSLAATTSPQLACGPTVHOCSENTYVKGCLFLGSLNRQPOK 120  
 DB 78 PLHPRPEAVNNSLGLSLAATTSPQLACGPTVHOCSENTYVKGCLFLGSLNRQPOK 136  
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRPMKELVSTIMEQLKSKTLFSLMQYSEEF 180  
 DB 137 VPDATPECPHQEMDIVFLIDGSGSIQDQDNQMGFVQAVMGQFEGDTLFLALMQYSNLL 196  
 QY 181 RIHFTFKEQNENPESLKIPIQLLGRTHATGLKRVRELNITNGARKNAFKLFL 240  
 DB 197 KIHFTFQRTSPSQSLVDPIVLQGLTFTATGILTVTLFHHKNGARKSAKILIVI 256  
 QY 241 TDGEKFGDPLGVEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNTVASKPPDDHVFQAN 300  
 DB 257 TDQOKYKDPLESDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNTVASKPPDDHVFQAN 316  
 QY 301 NFEALKTQVQREKIFAJEGTQSGSSSFHEMSQEGFSAITSNGPLISTVGSYDWAG 360  
 DB 317 NFEALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMQGLFLGAVGSFWSG 376  
 QY 361 GVPLYTSKSKSTFINNTRVDSMDNAYLGYAAAIILNRVQSVLVLGAPRYQHTGLVAMFR 420  
 DB 377 GAFLYPPNMSPTFINNTRVDSMDNAYLGYAAAIILNRVQSVLVLGAPRYQHTGLVAMFR 436  
 QY 421 QNTGMWSENVKGTQIGAYFGASLCSDVDNSGSDTLVLIGAPHYYEOTRGQNSVCL 480  
 DB 437 QVSRQWRKKAETVGTQIGAYFGASLCSDVDNSGSDTLVLIGAPHYYEOTRGQNSVCL 496



```
Qy 481 PRGQARWQCDAVLVYGGQGWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQVQWQCDAVLRGEGQHPGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSGTSPSHSQRISAKSLSPRLQYFGQSLGGODLTMDGLVDLTVCAGQCHVILLRSQ 600
Db 557 HGASEGSGTSPSHSQRISAKSLSPRLQYFGQSLGGODLTMDGLVDLTVCAGQCHVILLRSQ 616
Qy 601 PVLKRVKALMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLKRVKAVRSPVEKAVYRCWEKESALEAGDATVCLTIQKSSLDQL--GDIQSSVR 674
Qy 661 YDLALDSGRPHSRVFNTRKSTRQTVGLGTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 675 FDLALDPGLTSRAIFNETKNTLRRKTLGLGHCETLKLQLPNCIEDPSPVILRLNF 734
Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPPEKNGCNDNICODDLSITFSFMSLDCLVVG 780
Db 735 SLVREPISPNQRLRPVLAVGSDQLFTASLPPEKNGCQGLCEGDLGVTLSPGLQTLTVG 794
Qy 781 GPREENVTVVRNDEGDSYRTQVTFPPFLDLISYKRVSTLQNRQSRWRLACESASSTEV 840
Db 795 SSLELNVITVWNAGDSYGTVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 853
Qy 841 SGALKSTSCSINHPTPPENSEVFNITPDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 854 EG-LRSSRCVNHPIFHEGNGTFTVTFDVSFKATLGDRLMLMRASSENKAKSSKATF 912
Qy 901 QLELPVKAVVMVTVSHGVSTKYNF-TASENTSRVMQHOYVSNLQGRSLPISLVFLVP 959
Db 913 QLELPVKAVYVTVISQREESTKYPATSDKMKKEAERYVNVNLSQDLAISINFWVP 972
Qy 960 VRLNQTIVIMDRPQVTFSENLSCTCHKRLPSHSDPLAELRKAPVNCVSIACVQRIQCDI 1019
Db 973 VLLNGVAVWVMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQPCDV 1030
Qy 1020 PFGTQEPENATLKNLSPDWIKYSHNHLILVSTAEIILFNDSVFTLLPGQAGVRSQTE 1079
Db 1031 PSFSVQEELDFTLKNLSPGWRETLLQKKLVVSVAEITFDTSVSOLFQGAFFMAQME 1090
Qy 1080 TKVEPEVNPPLVLVSSVGLLILALITALYKLGFFKRYKQKMMSE 1128
Db 1091 MVLEDEVNATPIINGSSVGALLLALITALYKLGFFKRYKMELED 1139
RESULT 7
Q6KAS4
ID Q6KAS4 PRELIMINARY; PRT; 1188 AA.
AC Q6KAS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN Name=MFLJ00114;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:167-180(2004).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC EMBL; AK131133; BAD21383.1; -
DR GO; GO:0009897; C:external side of plasma membrane; IDA.
DR InterPro; IPR000413; Integrin_alpha.
```

```
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VMPADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR PROSITE; PS02334; VFWA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON TER 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.2%; Score 3306.5; DB 2; Length 1188;
Best Local Similarity 56.4%; Pred. No. 8.6e-211;
Matches 643; Conservative 172; Mismatches 306; Indels 19; Gaps 7;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGLYOCDSYSGSCPEI 60
Db 39 FNLDAAEKPTHEHMDGAERFHSVLQYDSSWVVGAPKEIKATNQIGGLYKCGYHTGNCPEI 98
Qy 61 RLQVPEAVNMSLGLSLAATTPPOLLIACGPTVHOTCGENTVVKGLCFGLFGLNLRQPOK 120
Db 99 SLQVPEAVNMSLGLSLAATNPSSLACGPTVHHTCRENIYLTGLCFLLSSFFQS-QN 157
Qy 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMPEQLKKSKTLFLSLMQVSEEF 180
Db 158 PTAQCEPKQDDIVFLIDGSGSIISSTDFEKMDFVKAQMSQLQRPSTFSLMQFSDYF 217
Qy 181 RIHFTFKFONNPNSRIKPTQLLGRTHATGRKVVRELPNITNGARKNAKLLFLL 240
Db 218 RVHFTFNFISSPLSLDLSVRQLRGVYTYASAIAKHVITELFTTQSGARQATKVLVI 277
Qy 241 TDGEKFGDPLGVEDVPELDREGVIRVILGFCDAFSEKSRQELNATVASKPPRDHVFQAN 300
Db 278 TDGRKQGNLSDYSVIPMAEASIIRYAIGVGAFAFNEHSHKELKAIASMPSEHYVFSVE 337
Qy 301 NFEALKTVQNLUREKIFATEGTQTSSESEHEMSQEGFSAATISNGPLLSTVSGYDWAQ 360
Db 338 NFDALNDIENKEKIFALEGTETSSSTFELEMSQEGFSAVTPDGPVLGAVGFSWSG 397
Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVILGAPRYOHIGLVAMPR 420
Db 398 GAFLYPSNMRPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVIET 457
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQSVWCPL 480
Db 458 QESRHWRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDLLVLIGVPHYEHTRGGQSVCCPM 517
Qy 481 PRGQARWQCDAVLVYGGQGWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 518 P-GVGRWFCGTLTGEGQHPGRFGAALTVLGVDVNGDSLADVAIGAPGEENRGAVYIF 576
Qy 541 HGTSGSGTSPSHSQRISAKSLSPRLQYFGQSLGGODLTMDGLVDLTVCAGQCHVILLRSQ 600
Db 577 HGASRQDIAPSPQSQRISAKSLSPRIQYFGQSLGGODLTMDGLVDLTVCAGQCHVILLRTR 636
Qy 601 PVLKRVKALMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 637 PILRVSTVHFTPEAISRSVFECQEQVAPQTLSDATVCLHIESPKQL--GDLRSTVT 694
Qy 661 YDLALDSGRPHSRVFNTRKSTRQTVGLGTQTCETLKLQLPNCIEDPSPVILRLNF 720
Db 695 FDLALDPGLTSRAIFNETKNTLRRKTLGLGHCESVKLLLPACVEDSVTPTTLRLNF 754
Qy 721 SLVGTPLSAFGLNLRPVLAEADQRLFTALFPPEKNGCNDNICODDLSITFSFMSLDCLVVG 780
Db 755 SLVGVPTSSLQNLQPLAVDDQTYFTASLPPEKNGGADHICQDDLSVVFPPDLKTLVVG 814
Qy 781 GPREENVTVVRNDEGDSYRTQVTFPPFLDLISYKRVKSTLQ-----NORSQSR 829
Db 815 SDLELNVITVSRNDEGDSYGTITVTLFYPVGLSFRRVAEQGVFLRKKEDQQWRGQSLH 874
```



RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,  
 RA Tsuchiya H.;  
 RT "Isolation of genes selectively expressed by dendritic cells.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 CC !- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It  
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell  
 CC interaction during inflammatory responses. It is especially  
 CC important in monocyte adhesion and chemotaxis (By similarity).  
 CC !- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X  
 CC associates with beta-2 (By similarity).  
 CC !- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC !- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC !- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC !- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC !- SIMILARITY: Contains 1 VWFA domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 DR EMBL: AF211864; AAF23492.1; --  
 DR HSSP: P20702; IN3Y.  
 DR MGD: MGI:96609; Itgax.  
 DR GO: GO:0009897; C:external side of plasma membrane; IDA.  
 DR InterPro: IPR000413; Integrin\_alpha.  
 DR InterPro: IPR002035; VWF\_A.  
 DR Pfam: PF01839; FG-GAP; 3.  
 DR Pfam: PF00357; Integrin\_alpha; 1.  
 DR Pfam: PF00092; VWF; 1.  
 DR PRINTS: PR01185; INTEGRINA.  
 DR PRINTS: PR00453; VWFADOMAIN.  
 DR SMART: SM00191; Int\_alpha; 5.  
 DR SMART: SM00327; VWF; 1.  
 DR PROSITE: PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE: PS0234; VWFA; 1.  
 DR Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 19 Potential.  
 FT CHAIN 20 1169 Integrin alpha-X.  
 FT DOMAIN 20 1116 Extracellular (Potential).  
 FT TRANSMEM 1117 1137 Potential.  
 FT DOMAIN 1138 1169 Cytoplasmic (Potential).  
 FT REPEAT 34 87 FG-GAP 1.  
 FT REPEAT 88 7 FG-GAP 2.  
 FT DOMAIN 152 330 VWFA.  
 FT REPEAT 402 402 FG-GAP 3.  
 FT REPEAT 403 454 FG-GAP 4.  
 FT REPEAT 456 518 FG-GAP 5.  
 FT REPEAT 519 577 FG-GAP 6.  
 FT REPEAT 582 634 FG-GAP 7.  
 FT CA\_BIND 467 475 Potential.  
 FT CA\_BIND 531 539 Potential.  
 FT CA\_BIND 594 602 Potential.  
 FT SITE 1140 1144 GFFKR motif.  
 FT DISULFID 69 76 By similarity.  
 FT DISULFID 108 126 By similarity.  
 FT DISULFID 656 711 By similarity.  
 FT DISULFID 770 776 By similarity.  
 FT DISULFID 858 873 By similarity.  
 FT DISULFID 1007 1031 By similarity.  
 FT DISULFID 1036 1041 By similarity.  
 FT CARBOHYD 89 89 N-linked (GlcNAc. . .)  
 FT CARBOHYD 267 267 N-linked (GlcNAc. . .) (potential).  
 FT CARBOHYD 393 393 N-linked (GlcNAc. . .) (potential).  
 FT CARBOHYD 734 734 N-linked (GlcNAc. . .) (potential).  
 FT CARBOHYD 949 949 N-linked (GlcNAc. . .) (potential).  
 FT CARBOHYD 1059 1059 N-linked (GlcNAc. . .) (potential).  
 FT CARBOHYD 1084 1084 N-linked (GlcNAc. . .) (potential).

SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;  
 Query Match 56.1%; Score 3297.5; DB 1; Length 1169;  
 Best Local Similarity 56.2%; Pred No. 3.3e-210;  
 Matches 641; Conservative 174; Mismatches 306; Indels 19; Gaps 7;  
 1 FNLDTENAMTFQENARGGQSVVQLQGSRRVVVVGAPQRIEVAANQRGSLVQCDYSTGSCBPI 60  
 20 FNLDAAELKTHFMDGAEFGHSLVQYDSSVWVVGAPKEIKATNQIGGLYKCGYHTGNCBPI 79  
 61 RLQVPEAVNMSLGLSLAATSPPLLACGPTVHTCTSENTVVKGLGCLFGLNLRQOPK 120  
 80 SLQVPEAVNMSLGLSLAATNPSSLACGPTVHTCTSENTVVKGLGCLFGLNLRQOPK 138  
 121 PEARLGGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEQLKSKTFLSFMQYSEBF 180  
 139 FTAQOECPKQDDIVFLIDGSGSISSDTEKMLDFVKAQMSQLQSPSTRFSLMQESDYF 198  
 181 RIHFTKFEQNNPNRSLIKPITQLLGRTHATGRKVVRELFNTNGARKNAKFLFL 240  
 199 RVHFTFNNFISTSPSLGSLVRLRGVYTTASAKHVITELFTTQSGARQDATKVLVI 258  
 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAERSEKSRQELANTVASKPRDHVFOAN 300  
 259 TDRKQGNLSDVSIPIVMAEASIIRYAIGVGKAFYNEHSHKQELKAIAMPSHEVYFVSE 318  
 301 NFEALKTVONQLREKIFAIEGTQGTSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAG 360  
 319 NFDALKDIENQKEKIFAIEGTETSSSTFELEMSQEGFSAVTPDPVGLGAVGSPWSG 378  
 361 GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAAILRNVRQSLVGLCAPYQHIGLVAMPR 420  
 379 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTGKVVIET 438  
 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVETRGQGVSCPL 480  
 439 QESRHWPKSEVRGTQIGSYFGASLCSVDMDRGGSTDVLVIGAPHYVETRGQGVSCPM 498  
 481 PRGQARWOCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYLF 540  
 499 P-GVGRWEGGTTLHGEQHPWGRFGAALTVLGDVNGDSLADVAIGAPGEENRGAYIF 557  
 541 HTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
 558 HGASRQDIAPSPSORISASQIPRIQYFGQSLSGGGQDLTMDGLVDLAVSGSKRVLRLTR 617  
 601 PVLRVKAIEMFPREVARNVFECNDQVVKGEAGEVRCVCHVCKSTRDLREGIOISVVT 660  
 618 PILRVSPVTHFTPAEISRSVFECQEQVAPEQTLSDATVCLHIESPTQL--GDLRSTVI 675  
 661 YDLALDSGRPHSAVENETKNSRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
 676 FDLALDHGRLSTRAIFKEKTRALTRVKGILGNKHCSVKLLLPACVEDSVTPIRLNF 735  
 721 SLVGTPLSAFNGRLPVLAEADAQLFTALPFFPKKNCNDNICQDDLITITFSMLDCLVVG 780  
 736 SLVGVPISSILQNLQPLAVDDQYFTASLPFFKXCGADHICQDDLVSFVFGFPLKTLVVG 795  
 781 GPREFNVTVVRNDGEDSVRTQVTFPFLDLSYEKYSTLQ-----NORSORSWR 829  
 796 SDLELVNDVTVSNDGEDSVGTVTTLFPVGLSFRRAVGQVFLRKEDQWQRGGQSLH 855  
 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSE 889  
 856 LMCD--STPDRSQGLMSTSCSRHVIIFRGSSQMTFLVTFVSPKAEGLDRLLLRARVGE 913  
 890 NNMPRNKTEFQLELPVKYAVYVMVTVSHGVSTKYLNFTASE-NTSRVWQHYQVSNLQOR 948  
 914 NNVEGTPKTTFQLELPVKYAVYVMVTVSHGVSTKYLNFTASE-NTSRVWQHYQVSNLQOR 973  
 949 SLPLSLVFLVPLVRLNQTVIWRDQVTFSENLSCTCHKRLPSHSDFLAELRKAPVNCVS 1008  
 974 DVPVSINFVWPIELKGEAVW-TVMVSHQPQPLTQCYNRLKLTQTFDILLTHMQKSPVLDGS 1032

```

QY 1009 IAVCORIQDIPFPGIQEENATLKNLSFDWYIKTSHNHLIVSTAILENDSVFTLLP 1068
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1033 IADCHLKDIPSLGILDELIFILKGNLSFGWISQLOKKVLLSEAEITNTSVYSQLP 1092
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 1069 GQAFVRSQTETKVFPEFVNPPLPIVGVSSVGGLLALLIITAALYKLGFFKQYKDMNSE 1128
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 1093 GQAFVRSQTETKVFPEFVNPPLPIVGVSSVGGLLALLIITAALYKLGFFKQYKDMNSE 1152
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 10
ID ITAD RAT STANDARD; PRT; 1161 AA.
AC Q90VE7;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien W.M.; VanderVieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT "Cloning of rat alpha D, a novel beta 2 integrin.";
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -! FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC VCAM1. May play a role in the atherosclerotic process such as
CC clearing lipoproteins from plaques and in phagocytosis of blood-
CC borne pathogens, particulate matter, and senescent erythrocytes
CC from the blood (By similarity).
CC -! SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC associates with beta-2 (By similarity).
CC -! SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -! DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -! SIMILARITY: Belongs to the integrin alpha chain family.
CC -! SIMILARITY: Contains 7 FG-GAP repeats.
CC -! SIMILARITY: Contains 1 VWFA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWFA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS0234; VWFA; 1.
DR KEGG; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
DR Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT DOMAIN 20 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Potential.
FT DOMAIN 1122 1161 Cytoplasmic (Potential).
FT REPEAT 34 87 FG-GAP 1.

```

```

FT REPEAT 88 ?
FT DOMAIN 152 334 FG-GAP 2.
FT REPEAT 352 402 VWFA.
FT REPEAT 403 454 FG-GAP 3.
FT REPEAT 455 517 FG-GAP 4.
FT REPEAT 518 577 FG-GAP 5.
FT REPEAT 578 634 FG-GAP 6.
FT REPEAT 635 697 FG-GAP 7.
FT CA_BIND 467 475 Potential.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GFFKR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 54.9%; Score 3225.5; DB 1; Length 1161;
Best Local Similarity 57.2%; Pred. No. 2.1e-205;
Matches 645; Conservative 163; Mismatches 306; Indels 13; Gaps 9;

QY 2 NLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRQSLVQCDYSTGSCCEPIR 61
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 21 NLDVEEPIVFEADAASFGQTVVQFGSRLVVGAPLEAVAVNQTGLYDCAATGMCQPIV 80
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 62 LQVPEAVNMSLGLSLAATTSPPOLLACGPTVHCTSENTVVKGLCFGLGNLRQPOKPF 121
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 81 LRSPLAVNMSLGLSLVATNNAQLACGPTAQRACVKNMYAKGSCLLIGSSL-QFIQAV 139
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 122 PEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLSLMOYSEFR 181
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 140 PASPECPQEMDIAFLIDGSGSINQDFAQMKDFVXALMGFEFASTTLFSLMOYSNLK 199
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 182 IHFTKFEQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNTNGARKNAKILFLTL 241
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 200 THFTTFEKNILDPSLVDPVQLQGLTYATGIRTWEEELFHSKNGSRKSAKILLVIT 259
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 242 DGEKFGDPLGYEDVPELDREGVTRYVVGFDAPFRSEKSRQELNTVASKPRDHVFOANN 301
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 260 DGQYRDPLEYSVDVIPAADKAGIIRYAGVGDAFQEPATKELNLTIGSAPPQDHVFKVGN 319
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 302 FEALKTQVQNLREKIFALEGTGTGSSSFHEHMSOEGFSAAITSNGLPLSTVSGVDWAGG 361
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 320 FAALRSIQRLQEKIFALEGTGTGSSSFHEHMSOEGFSAAITSNGLPLSTVSGVDWAGG 379
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 362 VFLYTSKEKSTFFINWTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMPQ 421
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 380 AFLYPNTRPTFFINNSQENVDMRDSYLGYSTAVAFWKGVHSLILGAPRHQHTGKVITQ 439
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 422 NTGMWESNANVKGTQIGAYFGASLCSVDNSNGSTDVLVIGAPHYEQTRGGQVSVCPPLP 481
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDNRDGSDDLVLGAPHYEQTRGGQVSVFPVP 499
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 482 RQQRARWQCDVLYGEGQWPGFAGALTVDGVNGDKLTDVAICAGEEDNRGAVLHF 541
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 500 -GVGRWQCEATLHGEQHPWGRFGVALTVLDGVNGDNLDVAITGAFCEESRGAVIFH 558
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
QY 542 GTSGSGISPSHSQRISAGSKLSPRLQYFQSLSGQDLTMDGLVLTVAQGHVLLLSQP 601
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 559 GASPLEIMPSPQSVTGSQSLRLQYFQSLSGQDLTQDGLVLDVLAQGHVLLLSLP 618
|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

```
QY 602 VLRVKAIMFNPVARNVFECDQVVKKEAGEVVRVCLHVOKSTDRDLRBEQIQSVVTV 661
DB 619 LKVLBSIRFAEMVAKAVQWERTPTVLEAGEATVCLVHKGSPDLL--GNVGSVRY 676
QY 662 DLALDSGRPHSAVNETKNSRRQTQVLGLTQTCTETLKQLQPCNIEOPVSPVLRILNFS 721
DB 677 DLALDPGLISRAIDFENKNCVLTGRKTLGLGDHCTETVKLLPDCVEDAVSPLILNFS 736
QY 722 LVGTPLSAPGNRPLVAEDAQLFTALPFFKNCNGNDNICQDDLITTSFMSLDCILVGG 781
DB 737 LVRDSASP--RNLPVLAVGSQDHITASLPFFKNCQELLCEGLGISEFNSGLQVLVGG 795
QY 782 PFEFNVTVTRNDGDSYRTQVTFPPLDLSYRKVSTLQNRQSRLACASASTVS 841
DB 796 SPELTVTVTVWMEGDSYGLTKFYFAGLSYRRVTGTQ-QPHQVPLRLACAEFAAQSD 854
QY 842 GALKSTCSINHIPIFENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTBFQ 901
DB 855 --LRSSCSINHIPIFEGAKTTFMITFDVSYKAFGLGRLLLRAKASSENKPDTKTAPQ 912
QY 902 LELPKYAVYVMTVSHGVSTKYNFTASNTGR-VWQHOYQVSNLQGRSLPISLFLVPV 960
DB 913 LELPKYVTVYTLISROEDSTNHFSSHGRRQEAHRYRVNLSPLKLAVRVFWVPV 972
QY 961 RLNQIVINDRPOVTFSENLS--TCHTKERLPSHSDFLAELRKAPVNCVSIACQRCODI 1019
DB 973 LINGVAVMD---VTLSSPAQGVSCVSKPPQNPDLTQIQRSLVDCSADCLHFRCDI 1029
QY 1020 PFEQIEEENATLKNGLSDWIKTSHNHLIVSTAELFNDVFTLLPGQAFVRSQTE 1079
DB 1030 PSLDIQDELDFILRNLNLSFGWVSQTLQEKVLLVSEAEITFDTSVYSQLPQAEFLRAQVE 1089
QY 1080 TKVEFEVNPVPLVGVSSVGLLLALITALYKLGFFKRYQKMM 1126
DB 1090 TTLEEVYVVEPIFLVAGSSVGLLLALITVLYKLGFFKRYQKML 1136

RESULT 11
ITAL_BOVIN STANDARD; PRT; 1165 AA.
AC P61625;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-L precursor (leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=ITGAL;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OC NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14697514; DOI=10.1016/j.gene.2003.09.043;
RA Felt T., Zecchin L., Baise E., Desmecht D.;
RT "The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,
RT characterisation and comparison with the human and murine
RT glycoproteins."
RL Gene 325:97-101(2004).
CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC mediated killing, and antibody dependent killing by granulocytes
CC and monocytes (By similarity).
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC associates with beta-2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- DOMAIN: The integrin I-domain (inset) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
```

```
CC CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC CC -!- SIMILARITY: Contains 1 VWFA domain.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; AY267467; AAP94035.1; -.
CC CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC CC PROSITE; PS0234; VWFA; 1.
CC CC Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
CC Repeat; Signal; transmembrane.
CC CC SIGNAL 1 23 Potential.
CC CC CHAIN 24 1165 Integrin alpha-L.
CC CC DOMAIN 24 1084 Extracellular (Potential).
CC CC TRANSMEM 1085 1105 Potential.
CC CC DOMAIN 1106 1165 Cytoplasmic (Potential).
CC CC REPEAT 40 88 FG-GAP 1.
CC CC REPEAT 89 146 FG-GAP 2.
CC CC REPEAT 153 324 VWFA.
CC CC REPEAT 347 397 FG-GAP 3.
CC CC REPEAT 398 452 FG-GAP 4.
CC CC REPEAT 454 513 FG-GAP 5.
CC CC REPEAT 515 572 FG-GAP 6.
CC CC REPEAT 575 627 FG-GAP 7.
CC CC CA_BIND 465 473 Potential.
CC CC CA_BIND 527 535 Potential.
CC CC CA_BIND 587 595 Potential.
CC CC SITE 1111 1115 GPFKR motif.
CC CC DISULFID 71 78 By similarity.
CC CC DISULFID 108 126 By similarity.
CC CC DISULFID 650 704 By similarity.
CC CC DISULFID 768 774 By similarity.
CC CC DISULFID 842 858 By similarity.
CC CC DISULFID 994 1009 By similarity.
CC CC DISULFID 1017 1048 By similarity.
CC CC CARBOHYD 33 33 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 86 86 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 185 185 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 646 646 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 667 667 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 723 723 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 859 859 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 894 894 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 929 929 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 1056 1056 N-linked (GlcNAc..) (Potential).
CC CC CARBOHYD 1067 1067 N-linked (GlcNAc..) (Potential).
CC CC SEQUENCE 1165 AA; 128725 MW; DAEBA3AF1E1463CB CRC64;
CC CC -----
CC CC Query Match 26.6%; Score 1561; DB 1; Length 1165;
CC CC Best Local Similarity 34.4%; Pred. No. 1.4e-94;
CC CC Matches 407; Conservative 200; Mismatches 468; Indels 108; Gaps 33;
CC CC -----
QY 1 FNLDTENAMTFQENARG--FGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCE 58
DB 24 YNLDRVHVQNFSPFPLAGRHFGYRVQLQV-GNGVVVGAPSE---GNSMGNYQQCPETGDC 79
QY 59 PIRLOQVEAVNMSLGLSLAAATSPPLQLACGPTVHQTGSENTYVKGLCFGLFGSNLRQOP 118
DB 80 PVTLS--SNYTSKYLQMTLATPSTDNLACDPGLSRTCDQNYLSGLCYLIHENLRGPV 137
QY 119 QKFPEALRCPOEDSDIAFLIDSGSIIIPDFFRMKELVSTIMEQLKKSKTLPSLMOYSE 178
DB 138 LQGHPCGYQECIKGNVDLVFLFDGNSLQODEFEKIVDFMKDVMKKLSNSYSQFAAVQFST 197
QY 179 ERLIHPTKEFQNNPNRSLIKPIPTOLLGRTHATGLRKVVRLEFNI TWGARKNAFKILF 238
DB 198 YRTEFTFLDIYRQKDPDALLAGVHKMLRLTWTFTFGAINVAKVEVFPDGLGARPDATKVI 257
```





QY	652	EQIOSVVTDIALDSGRPHSAVNEKSTRROTQVLGLTQTCETLKLQLPNCIEDPV	711
Db	662	QGHVLANITYTLQLDGHTRSGLPFGGKHKLIIGNTAVTPV-KSCFVFHFHFPICIQDLI	720
QY	712	SPIVLRNFSL---VGTPLS--AFGNLRPVLAEDAQRFTALFPPEKNCNNDNICODDLS	766
Db	721	SPINVLSYSLWEEGTTPDRALDRDIPPIKPSPHLETKEIPPEKNGCEDKNCADLK	780
QY	767	ITFSFMSLDCLVVGGRPREFNVTVRNDCGESYRTQVTFPPFLDLSYRKVSTLQNRQSR	826
Db	781	LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWQVTLSPFOGLSFRKVEIL---KPHS	837
QY	827	SNRLACESASTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANV	886
Db	838	HVPVGCCELPEEAVVHS-RALSCNVSPIFGEDSMVDIQVMFNTLQKSGWGFIELQANV	896
QY	887	TS-----ENNMPRTNKTETQLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQH	938
Db	897	SCNNEDSSILLEDNSATTS-----IPVMYPINVLTKQOENSTLYISFTPKSPRIHHVKH	949
QY	939	QYQV-----SNLQGRSLPISLFLVPRLNQTVI---WD---RPQVTFSS-ENLSSTCHTKE	987
Db	950	IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMBPPVNCSPRNLESPSDEAE	1008
QY	988	RLPSSHDFLAELRKAPVNVCSIAVCQRIODIPFGIOEENFATLKGNSLFDWYIKTSN	1047
Db	1009	-----SCSFGT--EPRCPIDF---ROEILVQVNGMVELRGTTIKAS-S	1044
QY	1048	HLIIVSTABILFNDSVFTLLPGOGAFVRSGTETKVEPFVFPNPLPLIVGSSVGGLLLLAL	1107
Db	1045	MLSLSLSSLAISFNSSKHFLHGRNASM-AQVVMKVDLVYEKMLYLVLSGIGLLLLFL	1103
QY	1108	ITAALYKLGFFKRYQKQKMM-----SEGPPGABPQ	1137
Db	1104	IFIALYKVGFFKRNLEKMEANVDASSEIPGEDAGQPELEKE	1145
RESULT 13			
AAQ90015			
ID	AAQ90015	PRELIMINARY;	PRT; 1166 AA.
AC	AAQ90015;		
DT	02-MAR-2004	(T-EMBLrel. 27, Created)	
DT	02-MAR-2004	(T-EMBLrel. 27, Last sequence update)	
DT	02-MAR-2004	(T-EMBLrel. 27, Last annotation update)	
DE	Lymphocyte function-associated antigen 1 alpha subunit CD11a.		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;		
OC	Bovidae; Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RA	Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;		
RT	"Molecular cloning and sequencing of bovine CD11a."		
RL	Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY382558; AAQ90015.1;		
SQ	SEQUENCE 1166 AA; 128723 MW; 758D8AFBA896C9DF CRC64;		
Query Match 26.2%; Score 1539; DB 2; Length 1166;			
Best Local Similarity 34.3%; Pred. No. 4e-93;			
Matches 405; Conservative 199; Mismatches 472; Indels 106; Gaps 33;			
QY	1	FNLDTENAMTFQENARG--FGQSVVLQSGRVVVGAPQELIVANQSRSLYQCDYSTGCE	58
Db	25	YNLDRVHVQNFSPFLAGRHFGYRLQV-GNGVVVGAPSE---GNSMGNYQCQPETGDCL	80
QY	59	PIRLQVPVEAVNMSLGLSLAATSPPOLLAGCTVHTQCSNTYVXGLCLFLGFSNLRQOP	118
Db	81	PVTLS--SNYTSKYLGNLTATDPTSDNLLACDGLSRTCDQNIYLSGLCYLIHENLRGPV	138
QY	119	QKPEALRGCPQEDSDIAFLIDGSGIIPHDFRMKELVSTIMEQKKSKTLFSLMOYSE	178

RESULT 14  
ITAL\_HUMAN



```

ID AC P20701.043746; STANDARD; PRT; 1170 AA.
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=ITGAL; Synonyms=CD11A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX MEDLINE=89139587; PubMed=2537322;
RA Larson R.S., Corbi A.L., Berman L., Springer T.;
RT "Primary structure of the leukocyte function-associated molecule-1
RT alpha subunit: an integrin with an embedded domain defining a protein
RT superfamily.";
RL J. Cell Biol. 108:703-712(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX MEDLINE=96036067; PubMed=7479767;
RA Qu A., Leahy D.J.;
RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L
RT beta 2) integrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX MEDLINE=96398682; PubMed=8805579;
RA Qu A., Leahy D.J.;
RT "The role of the divalent cation in the structure of the I domain from
RT the CD11a/CD18 integrin.";
RL Structure 4:931-942(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX MEDLINE=99425288; PubMed=10493852;
RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA Cottens S., Weitz-Schmidt G., Hommel U.;
RT "Structural basis for LFA-1 inhibition upon lovastatin binding to the
RT CD11a I-domain.";
RL J. Mol. Biol. 292:1-9(1999).
CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC mediated killing, and antibody dependent killing by granulocytes
CC and monocytes.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC associates with beta-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P20701-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P20701-2; Sequence=VSP_002738;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Leukocytes.
CC -!- DOMAIN: The integrin I-domain (insert) is a VFMA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC CC -!- SIMILARITY: Contains 1 VWFA domain.
CC CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;
CC CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11a.htm".
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@sib-sib.ch).
CC CC -----
CC CC EMBL; Y00796; CAA68747.1; -.
CC CC EMBL; AC002310; AAC31672.1; -.
CC CC PIR; S03308; S03308.
CC CC PDB; 1CQP; X-ray; A/B=153-334.
CC CC PDB; 1DGO; NMR; A=149-336.
CC CC PDB; 1LFA; X-ray; A/B= -.
CC CC PDB; 1MJN; X-ray; A=153-331.
CC CC PDB; 1MQ8; X-ray; B/D=155-331.
CC CC PDB; 1MQ9; X-ray; A=152-330.
CC CC PDB; 1MQA; X-ray; A=152-330.
CC CC PDB; 1ZON; X-ray; @=150-336.
CC CC PDB; 1ZOO; X-ray; A/B=150-336.
CC CC PDB; 1ZOP; X-ray; A/B=150-336.
CC CC Genew; HGNC:6148; ITGAL.
CC CC MIN; 153370; -.
CC CC GO; GO:0008305; C:integrin complex; TAS.
CC CC GO; GO:0006928; P:cell motility; TAS.
CC CC InterPro; IPR000413; Integrin_alpha.
CC CC InterPro; IPR002035; VWF_A.
CC CC Pfam; PF01839; FG-GAP; 3.
CC CC Pfam; PF00357; Integrin_alpha; 1.
CC CC Pfam; PF00092; VWFA; 1.
CC CC PRINTS; PR01185; INTEGRINA.
CC CC PRINTS; PR00453; VWFADOMAIN.
CC CC SMART; SM00191; Int_alpha; 5.
CC CC SMART; SM00327; VWFA; 1.
CC CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC CC PROSITE; PS00234; VWFA; 1.
CC CC 3D-structure; Alternative splicing; Calcium; Cell adhesion;
CC CC Direct protein sequencing; Glycoprotein; Integrin; Magnesium;
CC CC Receptor; Repeat; Signal; Transmembrane.
CC CC SIGNAL 1 25
CC CC CHAIN 26 1170 Integrin alpha-L.
CC CC DOMAIN 26 1090 Extracellular (Potential).
CC CC TRANSMEM 1091 1111 Potential.
CC CC DOMAIN 1112 1170 Cytoplasmic (Potential).
CC CC REPEAT 42 91 FG-GAP 1.
CC CC REPEAT 92 149 FG-GAP 2.
CC CC DOMAIN 156 327 VWFA.
CC CC REPEAT 350 400 FG-GAP 3.
CC CC REPEAT 401 455 FG-GAP 4.
CC CC REPEAT 457 516 FG-GAP 5.
CC CC REPEAT 518 575 FG-GAP 6.
CC CC REPEAT 578 630 FG-GAP 7.
CC CC CA_BIND 468 476 Potential.
CC CC CA_BIND 530 538 Potential.
CC CC CA_BIND 590 598 Potential.
CC CC SITE 1115 1119 GFFKR motif.
CC CC DISULFID 73 80 By similarity.
CC CC DISULFID 111 129 By similarity.
CC CC DISULFID 653 707 By similarity.
CC CC DISULFID 771 777 By similarity.
CC CC DISULFID 845 861 By similarity.
CC CC DISULFID 998 1013 By similarity.
CC CC DISULFID 1021 1052 By similarity.
CC CC CARBOHYD 65 65 N-linked (GlcNAc...) (Potential).
CC CC CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
CC CC CARBOHYD 188 188 N-linked (GlcNAc...) (Potential).
CC CC CARBOHYD 649 649 N-linked (GlcNAc...) (Potential).
CC CC CARBOHYD 670 670 N-linked (GlcNAc...) (Potential).
CC CC CARBOHYD 726 726 N-linked (GlcNAc...) (Potential).

```



alpha-subunit and its expression in COS cells.";  
 J. Immunol. 147:369-374(1991).  
 [2]  
 SEQUENCE OF 24-42.  
 MEDLINE=85188276; PubMed=3887182;  
 Springer T.A., Teplov D.B., Dreyer W.J.;  
 "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion  
 glycoproteins and unexpected relation to leukocyte interferon.";  
 Nature 314:540-542(1985).  
 CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,  
 ICAM3 and ICAM4. Is involved in a variety of immune phenomena  
 including leukocyte-endothelial cell interaction, cytotoxic T-cell  
 mediated killing, and antibody dependent killing by granulocytes  
 and monocytes. Mice expressing a null mutation of the alpha-L  
 subunit gene demonstrate impaired tumor rejection and impaired  
 leukocyte recruitment.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L  
 associates with beta-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Leukocytes.  
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M60778; AAA39426.1; -.  
 DR PIR; I56126; I56126.  
 DR HSSP; P20701; IDGQ.  
 DR MGP; MGI:96606; Itgal.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR02035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 1.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWFA; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS0234; VWFA; 1.  
 KW Calcium; Cell adhesion; Direct protein sequencing; Glycoprotein;  
 KW Integrin; Magnesium; Receptor; Repeat; Signal; Transmembrane.  
 FT SIGNAL 23  
 FT CHAIN 24  
 FT DOMAIN 24 1163 Integrin alpha-L.  
 FT TRANSMEM 24 1087 Extracellular (Potential).  
 FT DOMAIN 1088 1108 Potential.  
 FT DOMAIN 1109 1163 Cytoplasmic (Potential).  
 FT REPEAT 39 88 FG-GAP 1.  
 FT REPEAT 89 146 FG-GAP 2.  
 FT DOMAIN 153 325 VWFA.  
 FT REPEAT 348 398 FG-GAP 3.  
 FT REPEAT 399 454 FG-GAP 4.  
 FT REPEAT 455 514 FG-GAP 5.  
 FT REPEAT 516 573 FG-GAP 6.  
 FT REPEAT 576 628 FG-GAP 7.  
 FT CA\_BIND 466 474 Potential.  
 FT CA\_BIND 528 536 Potential.  
 FT CA\_BIND 588 596 Potential.  
 FT SITE 1111 1115 GFFKR motif.  
 FT SITE 70 77 By similarity.  
 FT DISULFID 108 126 By similarity.  
 FT DISULFID 147 199 By similarity.  
 FT DISULFID 651 705 By similarity.  
 FT DISULFID 767 773 By similarity.

FT	DISULFID	840	856	By similarity.
FT	DISULFID	993	1009	By similarity.
FT	DISULFID	1017	1048	By similarity.
FT	CARBOHYD	86	86	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	185	185	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	270	270	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	444	444	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	668	668	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	724	724	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	728	728	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	776	776	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	857	857	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	880	880	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	890	890	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	899	899	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	927	927	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1056	1056	N-linked (GlcNAc. . .) (Potential).
SQ	SEQUENCE	1163	AA; 128343	MM; A7A3078489E8232F CRC64;

Query Match 25.8%; Score 1516.5; DB 1; Length 1163;  
 Best Local Similarity 33.9%; Pred. No. 1.2e-91;  
 Matches 396; Conservative 216; Mismatches 466; Indels 89; Gaps 36;

QY	1	ENLDTENAMTFQENA-RGFGQSVVQLQSGRVVVGAPQELIVANQRGSLYQCDYSTGSCPE	59
Db	24	YNLDRTPQSLAQAGRHFGYQVLQIEDG-VVVGAPGE---GDNTGGLYHCRSTSEFCQP	79
QY	60	IRLQVPVEAVNMSLGLSLAATTPQLIACGPTVHQTCSENTYVKGCLFLFGSNLRQCPQ	119
Db	80	VSLH-GSNHSTSKYLGMTLATDAAGSLIACDPGLSRCTQNTYLSGLCYLPQSLGPM	138
QY	120	KPPALRCQPEDSDIAFLIDSGSIIPHDPFRMKELVSTIMEQLKSKTILFSLMOYSEE	179
Db	139	QNRPAYQECMKGVLDVFLFDGQSOLDRKDKILEFMKDVWRKLSNTSYQFAAVQFSTD	198
QY	180	FRIHFTKEF-QNNPNRESLIKPTOLLGRTHATGLRKVVRELFININGARKNAFKILF	238
Db	199	CRTEFTLDYVYKNNPDVLGSGVQPMFLNTTFRAINVYVAHVFKESGARPDAKVLV	258
QY	239	LLTDGEKF--GDPLGYEDVPELDEGVIRVVLGFGDAFRSEKSRQELNTVASKPRDHV	296
Db	259	IITDGEASDKGNISAHD-----ITRIIIGIKHFVSQVQKTLHIIFASEVEBFV	309
QY	297	FOANPEALKTVONOLREKIFAISGTQTGSSSEHEMSQBFSAATNSGPELLSTVGSY	356
Db	310	KILDTFEKLDLFTDLQRIYAIETGNRQDLTSFNMELSSSGISADLSKGHAVVGVAK	369
QY	357	DWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGAAA-IILNRVQSLVLGAPRYQHIG	414
Db	370	DWAGFLDLREDLQGFVGOEPLTSDVRGGYLGTYVAMWTSRSRPLAAGAPRYQHV	429
QY	415	LVAMFR--QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQDR	472
Db	430	QVLLFQAPAEAGRWNOQTKEGTQIGSYFGGELCSVLDQDGEALLIGAPLFGEGRG	489
QY	473	GQSVVCPLPRQARWOCDAVLYGEOGPWRFGAALTVLGDVNGDKLTDVAIGAPGED	532
Db	490	GRVFTY---QRRQSLFEMVSELQDGPYPLGRFGAAITALTIDINGDRITDVAVGAPLEE	545
QY	533	NRGAVLPHGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGGQDLTDGDLVDTVAQG	592
Db	546	-QGAVYIFNGKPG-GLSPQSPQRIQGAQVFGIRWFGRSIHGVKDLGGDLADVVVGAEG	603
QY	593	HVLLLRQPVLRKVAIMEFNPREVARNVFECNDQVVKGBAG-EVRYCLHVQKSTRDLR	651
Db	604	RVVVLSPRPVVDVVTLELSPFPEIPVHEVECSYAREEQKHGKVLKAFRIKPLTPQ--P	661
QY	652	EGQIQSVVTVDLALDSGRPHSRVFNFTKXSTRQTVLGLTQTCETLKLQPLNCIEDPV	711
Db	662	QGRLLANLSTYQLQDGHMRMRSGLFPDGSHELSTSTIP-DKSCLDLFFHFHFCIQDLI	720
QY	712	SPIVLRNLNFSLV---GTPLSAFGN-LRPVLAEADAQRLFTALFPFEKNCNGNDNTCQDLSI	767

[illegible]

Search completed: November 9, 2004, 12:17:27  
Job time : 142.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds  
(without alignments)  
3950.365 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_23Sep04:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5839	99.3	1153	2	AAW65090 Human Bet
2	5839	99.3	1153	3	AAB07360 Human CD1
3	5839	99.3	1153	5	AAU80252 Human int
4	5839	99.3	1153	5	ABG61469 Human Bet
5	5839	99.3	1153	5	AAO14428 Integrin
6	5839	99.3	1153	7	ADD25615 Binding d
7	5829	99.1	1153	2	AAO4136 Alpha sub
8	5823.5	99.1	1152	8	ADM99589 Human int
9	5816.5	98.9	1152	8	ADP12435 Protein e
10	3460	58.9	1163	8	ADP44061 Human CD1
11	3446	58.6	1163	8	ADN02004 Human inf
12	3446	58.6	1163	8	ADQ17510 Human sof
13	3437	58.5	1163	2	AAO7120 p150.95 a
14	3423	58.2	1163	2	AAW65091 Human Bet
15	3423	58.2	1163	3	AAO7361 Human CD1
16	3423	58.2	1163	5	ABG61470 Human Bet
17	3421	58.2	1163	6	ABU07406 Protein d
18	3421	58.2	1163	7	ADG32005 Human hom
19	3388	57.6	1161	2	AAO78166 Human bet
20	3388	57.6	1161	2	AAW23049 Human bet
21	3388	57.6	1161	2	AAW57491 Human bet
22	3388	57.6	1161	2	AAW65089 Human Bet
23	3388	57.6	1161	2	AAW72825 Human alp
24	3388	57.6	1161	2	AAW73342 Human alp
25	3388	57.6	1161	3	AAB07359 Human alp

## ALIGNMENTS

### RESULT 1

AAW65090  
ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

DT 28-SEP-1998 (first entry)

XX DE Human Beta-integrin CD11b subunit protein.

Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
lung inflammation; acute respiratory distress syndrome; CD11b subunit;  
rheumatoid arthritis.

OS Homo sapiens.

XX PN US5728533-A.

XX PD 17-MAR-1998.

XX PF 07-JUN-1995; 95US-00485618.

XX PR 23-DEC-1993; 93US-00173497.

XX PR 05-AUG-1994; 94US-00286889.

XX XX 21-DEC-1994; 94US-00362652.

XX PA (ICOS-) ICOS CORP.

XX PI Van Der Vieren M, Gallatin WM;

XX XX WPI; 1998-206565/18.

XX XX Screening assay for modulators of integrin binding - using immobilised or

labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.  
Example 5; Fig 1A-D; 106pp; English.

This sequence represents a human beta-integrin CD11b subunit which is used to describe a method for identifying compounds that modulate the interaction of the beta-integrin alpha-d subunit with a binding partner of alpha-d which involves contacting an alpha-d polypeptide with an alpha-d binding partner, one of which is immobilised and the other of which is labelled, in the presence of a test compound, and determining if the compound affects binding between the alpha-d polypeptide and alpha-d binding partner, where the alpha-d polypeptide is alpha-d or its fragment comprising the cytoplasmic, transmembrane or extracellular domain of alpha-d. Compounds that modulate alpha-d binding could be used to treat.

ABG61468 Human Bet  
Aaw23064 Human Bet  
Aaw65106 Human Bet  
Aaw72837 Human alp  
Aaw73343 Human alp  
Aab07376 Human alp  
ABG61485 Human Bet  
Aaw78169 Rat alpha  
Aaw23062 Rat beta  
Aaw60004 Rat alpha  
Aaw72824 Rat alpha  
Aab07374 Rat alpha  
ABG61483 Rat Beta2  
Aaw65104 Rat beta-  
Aaw73345 Rat alpha  
Aaw23061 Mouse bet  
Aaw60003 Mouse alp  
Aaw65103 Mouse bet  
Aaw72836 Mouse alp  
Aaw73347 Mouse alp

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
CC and rheumatoid arthritis  
XX  
SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 2; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGRVVVVGAEIIVAAANQKSLYQCDYSTGSCBPI 60  
Db 17 FNLDTENAMTFOENARGFGQSVVQLQGRVVVVGAEIIVAAANQKSLYQCDYSTGSCBPI 76  
QY 61 RLQVPEAVNMISLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGCLFLGSLNLRQOPK 120  
Db 77 RLQVPEAVNMISLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGCLFLGSLNLRQOPK 136  
QY 121 FPEARLGCPOEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTKLFLSLMOYSEEF 180  
Db 137 FPEARLGCPOEDSDIAFLIDGSGIIPHDPRMKELVSTIMEQLKSKTKLFLSLMOYSEEF 196  
QY 181 RIHFTFKFQNNPNSRLIKPTQLLGRTHATGRLKRVVRELFNITNGARKNAFKILFL 240  
Db 197 RIHFTFKFQNNPNSRLIKPTQLLGRTHATGRLKRVVRELFNITNGARKNAFKILV 256  
QY 241 TDGEKFGDPLGVEDVPELDRGVIRYVLGFGDAFRSEKSRQELNIVASKPRDHVQAN 300  
Db 257 TDGEKFGDPLGVEDVPELDRGVIRYVLGFGDAFRSEKSRQELNIVASKPRDHVQAN 316  
QY 301 NFEALKTVQNLQREKIFALEGTQTGSSSFEHEMSQEGFSAAITSGNPLLSVGVSDWAG 360  
Db 317 NFEALKTVQNLQREKIFALEGTQTGSSSFEHEMSQEGFSAAITSGNPLLSVGVSDWAG 376  
QY 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAILNRVQSLVGLGAPYQHIGLVAMPR 420  
Db 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAILNRVQSLVGLGAPYQHIGLVAMPR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSCPL 480  
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSCPL 496  
QY 481 PRGQARWQCDVLYGEOQOPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWQCDVLYGEOQOPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLLRQ 600  
Db 557 HTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLLRQ 616  
QY 601 PVLRYKAIMERNPREVARNVFECNDQVVKGEKAGEVRVCLHVQKSTRDRRLRGOIQSVVT 660  
Db 617 PVLRYKAIMERNPREVARNVFECNDQVVKGEKAGEVRVCLHVQKSTRDRRLRGOIQSVVT 676  
QY 661 YDLALDSGRPHRAVENETKNSRRTOTQVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720  
Db 677 YDLALDSGRPHRAVENETKNSRRTOTQVLGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 736  
QY 721 SLVGTPLSAFNLRLPVLAEADAQLFTALFPFEKNCNDNICODDLITFTSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLRLPVLAEADAQLFTALFPFEKNCNDNICODDLITFTSFMSLDCLVVG 796  
QY 781 GREFNVTTVNRDGEDSVRTQVTFPPFDLSVRKYSTLQNSQSRSLACASSTEV 840  
Db 797 GREFNVTTVNRDGEDSVRTQVTFPPFDLSVRKYSTLQNSQSRSLACASSTEV 856  
QY 841 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900  
Db 857 SGALKSTSCSINHPIFENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 916  
QY 901 QLELPVKYAVYMWVTHSGVSTKYLNFTASENTSRVMQHVQVSNLQGRSLPISLVLFPV 960  
Db 917 QLELPVKYAVYMWVTHSGVSTKYLNFTASENTSRVMQHVQVSNLQGRSLPISLVLFPV 976

QY 961 RLNOTVIWDRPOVTFSENLSTCHTKERLPSHSDFLAELRKPAPVNVCSIAVCQRIQCDDIP 1020  
Db 977 RLNOTVIWDRPOVTFSENLSTCHTKERLPSHSDFLAELRKPAPVNVCSIAVCQRIQCDDIP 1036  
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFAVRSQTET 1080  
Db 1037 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQGAFAVRSQTET 1096  
QY 1081 KVEPPEVENPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSSGGPPGAPQ 1137  
Db 1097 KVEPPEVENPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSSGGPPGAPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; protein; 1153 AA.

XX

AC AAB07360;

XX

DT 17-JAN-2001 (first entry)

XX

DE Human CD11b protein sequence.

XX

KW Human; macrophage infiltration inhibition; alpha\_d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11b.

XX

OS Homo sapiens.

XX

PN WO200029446-A1.

XX

PD 25-MAY-2000.

XX

PF 16-NOV-1999; 99WO-US027139.

XX

PR 16-NOV-1998; 98US-00193043.

XX

PR 08-JUL-1999; 99US-00350259.

XX

PA (ICOS-) ICOS CORP.

XX

PI Gallatin MW, Van Der Vieren M;

XX

XX WPI; 2000-387751/33.

DR

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit  
PT macrophage infiltration and reduce inflammation at central nervous system  
PT injury sites.

XX

XX Example 5; Fig 1; 270pp; English.

XX

CC Integrins are a class of membrane-associated molecules that participate  
CC in cellular adhesion. Integrins are made up of an alpha subunit and a  
CC beta subunit. One class of human integrins are restricted to expression  
CC in white blood cells and have a common beta2 subunit: the leukocyte  
CC integrins, leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins  
CC have an important role in immune and inflammatory responses. The present  
CC protein sequence is the human integrin alpha subunit CD11b. This sequence  
CC was used in an alignment to identify a novel beta2 integrin alpha  
CC subunit: alpha d (AAA60014 and AAB07359). The present sequence has  
CC approximately 60% identity to the protein sequence of alpha d. The  
CC alpha d gene and protein may be useful in therapy for diseases linked to  
CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple  
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory  
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency  
CC (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the  
CC inhibition of macrophage infiltration at the site of a central nervous  
CC system injury. The monoclonal antibodies can also be used to detect and  
CC diagnose Crohn's disease

XX

SQ Sequence 1153 AA;		Query Match		99.3%; Score 5839; DB 3; Length 1153;	
		Best Local Similarity		98.9%; Pred. No. 0;	
		Matches 1124; Conservative		8; Mismatches 5; Indels 0; Gaps 0;	
QY	1	FNLDTENAMTFOENARGGQSVVQLQGSRRVVVGAPEIIVAAQNRGSLYQCDYSGSCPEI	60		
Db	17	FNLDTENAMTFOENARGGQSVVQLQGSRRVVVGAPEIIVAAQNRGSLYQCDYSGSCPEI	76		
QY	61	RLQVPVEAVNMSLGLSLAATSPPOLLAGCPTVHOTCSENTYVYKGLCFLFGSNLRQOPK	120		
Db	77	RLQVPVEAVNMSLGLSLAATSPPOLLAGCPTVHOTCSENTYVYKGLCFLFGSNLRQOPK	136		
QY	121	FPEALRGCPGSDSIAFLIDGSGSTIIPHDFRMKELVSTIMEOLKSKTILFSLWQYSEEF	180		
Db	137	FPEALRGCPGSDSIAFLIDGSGSTIIPHDFRMKELVSTIMEOLKSKTILFSLWQYSEEF	196		
QY	181	RIHFTTFEFONNPNRSLIKBITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL	240		
Db	197	RIHFTTFEFONNPNRSLIKBITQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI	256		
QY	241	TDGKFGDPLGYEDVIBELDREGVIRYVVGDFRSEKSKOELNTVASKPPRDHVFQAN	300		
Db	257	TDGKFGDPLGYEDVIBELDREGVIRYVVGDFRSEKSKOELNTVASKPPRDHVFQAN	316		
QY	301	NFEALKTVONOLREKIIFAIEGTQTGSSSFHEMSOEGFSAITNSGPLLSTVGSYDWAG	360		
Db	317	NFEALKTVONOLREKIIFAIEGTQTGSSSFHEMSOEGFSAITNSGPLLSTVGSYDWAG	376		
QY	361	GVFLYTSKEKSTFINMTRVSDMDNDVILGYAAAIIILNRVQSLVILGAPRYOHIGLVAMFR	420		
Db	377	GVFLYTSKEKSTFINMTRVSDMDNDVILGYAAAIIILNRVQSLVILGAPRYOHIGLVAMFR	436		
QY	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSCPL	480		
Db	437	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEYQTRGGQVSCPL	496		
QY	481	PRGQARWQCDVILYGGQGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVILF	540		
Db	497	PRGQARWQCDVILYGGQGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEEDNRGAVILF	556		
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTWGAQGHVLLLRSQ	600		
Db	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTWGAQGHVLLLRSQ	616		
QY	601	PVLVRKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660		
Db	617	PVLVRKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	676		
QY	661	YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCEIDPVSPIVILNLF	720		
Db	677	YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCEIDPVSPIVILNLF	736		
QY	721	SLVGTPLSAFGRNLRPLVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFFSMSLDCLVWG	780		
Db	737	SLVGTPLSAFGRNLRPLVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFFSMSLDCLVWG	796		
QY	781	GPREFNVTVVRNDGDSYQTVTFPPDLISYRKVSTLQNRQSRQSWRLACESASSTEV	840		
Db	797	GPREFNVTVVRNDGDSYQTVTFPPDLISYRKVSTLQNRQSRQSWRLACESASSTEV	856		
QY	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKXANVTSENMPRTNKTEF	900		
Db	857	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKXANVTSENMPRTNKTEF	916		
QY	901	QLELPVKYAVYVMVTSHGVTSKYLNFTASENTSRVMQHQYQVSNLGRQSLPISLVFLVPV	960		
Db	917	QLELPVKYAVYVMVTSHGVTSKYLNFTASENTSRVMQHQYQVSNLGRQSLPISLVFLVPV	976		
QY	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP	1020		
Db	977	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIP	1036		

QY 1021 FFGIOEBENATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1080

Db 1037 FFGIOEBENATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1096

QY 1081 KVEPFEVPNPPLIVGSSVGGLLLLALITAAALYKLGFEKROYKOWMSGGPPGAPQ 1137

Db 1097 KVEPFEVPNPPLIVGSSVGGLLLLALITAAALYKLGFEKROYKOWMSGGPPGAPQ 1153

RESULT 3

AAU80252

ID AAU80252 standard; protein; 1153 AA.

XX AAU80252;

AC AAU80252;

XX 15-JUL-2002 (first entry)

DT Human integrin 1 alpha-M subunit protein.

XX Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;

XX inflammatory disease; autoimmune disorder; Crohn's disease;

KW human immunodeficiency virus; HIV; myocardial infarction;

KW Sjorgen's syndrome; rheumatoid arthritis.

XX Homo sapiens.

OS

XX Key Location/Qualifiers

PH Misc-difference 499. -500

FT /note= "Encoded by GGG CAG AGG"

FT WO200218583-A2.

XX 07-MAR-2002.

XX 31-AUG-2001; 2001WO-US027227.

XX 01-SEP-2000; 2000US-0229700P.

XX (BLOO-) CENT BLOOD RES INC.

XX Springer TA, Shimoaka M, Lu C;

XX WPI; 2002-382964/41.

XX N-PSDB; ABK50046.

XX Modified integrin-I or integrin I-like domain polypeptide useful as an

XX immunogen to produce antibodies specific to polypeptide, comprises a

XX disulfide bond such that polypeptide is stabilized in a desired

XX conformation.

XX Disclosure; Page 109-112; 112pp; English.

XX This invention relates to a modified integrin-I or integrin I-like domain

XX polypeptide comprising at least one disulfide bond so that the domain is

XX stabilised in a desired conformation. The polypeptide of the invention

XX may have antiinflammatory or immunosuppressive activities. The

XX polypeptides of the invention have an open conformation and are useful as

XX immunogens to produce antibodies that selectively bind to integrin I-

XX domain; and for identifying a modulator of integrin activity, or of

XX interaction of an integrin and a cognate ligand. The polypeptide of the

XX invention, or antibodies (preferably anti-LFA-1 antibody) is useful for

XX treating or preventing an integrin mediated disorder which is an

XX inflammatory or autoimmune disorder in a subject and for inhibiting the

XX binding of an integrin to a cognate ligand such as Crohn's disease,

XX nephritis; human immunodeficiency virus (HIV), myocardial infarction,

XX Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic

XX composition comprising the peptide of the invention is useful for

XX treating an integrin mediated disorder in a subject. The polypeptides

XX and/or active or antigenic fragments are useful as reagents for diagnosis

XX of integrin-mediated disorders. The present sequence represents the human

XX integrin-1 alpha-M protein subunit used to generate the mutant

XX polypeptides of the invention



XX	Sequence	1153 AA
QY	Query Match	99.3%; Score 5839; DB 5; Length 1153;
QY	Best Local Similarity	98.9%; Pred. No. 0;
QY	Matches 1124; Conservative	8; Mismatches 5; Indels 0; Gaps 0;
QY	1	PNLDTENAMTFOENARGFGQSQWQLOQSGRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60
Db	17	PNLDTENAMTFOENARGFGQSQWQLOQSGRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76
QY	61	RLOVPVEAVNMSLGLSLAATSPQQLACGTVHTQTCSENTYVYKGLCFLFGSNLRQOPQK 120
Db	77	RLOVPVEAVNMSLGLSLAATSPQQLACGTVHTQTCSENTYVYKGLCFLFGSNLRQOPQK 136
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKELYSTIMEQLKSKTLFSLMOYSEEF 180
Db	137	FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKELYSTIMEQLKSKTLFSLMOYSEEF 196
QY	181	RIHTFKFQONPNPRSLIKIPITOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db	197	RIHTFKFQONPNPRSLIKIPITOLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256
QY	241	TGKEKFGDPLGYEDVIPELDREGVIRYVGLFGDAPRSEKSKQELNTVASKPPRDHVFQAN 300
Db	257	TGKEKFGDPLGYEDVIPELDREGVIRYVGLFGDAPRSEKSKQELNTVASKPPRDHVFQAN 316
QY	301	NFEALKTVONQLREKIPFAIEGTQTGSSSFHEMSQEGFAAITSNGLLSTVGSYDWAG 360
Db	317	NFEALKTVONQLREKIPFAIEGTQTGSSSFHEMSQEGFAAITSNGLLSTVGSYDWAG 376
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRNVSQSLVGLGAPRYOHIGLVAMFR 420
Db	377	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRNVSQSLVGLGAPRYOHIGLVAMFR 436
QY	421	QNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTRGGQVSVCP 480
Db	437	QNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTRGGQVSVCP 496
QY	481	PRGQARWQCDVLYGEOGPWGRGAALTVLGDVNGDKLTDVAIGAPGEEDNRGANVYL 540
Db	497	PRGQARWQCDVLYGEOGPWGRGAALTVLGDVNGDKLTDVAIGAPGEEDNRGANVYL 556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTGAGQHVLLLRQ 600
Db	557	HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTGAGQHVLLLRQ 616
QY	601	PVLRVKAIMEFNPREVARNFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db	617	PVLRVKAIMEFNPREVARNFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676
QY	661	YDLALDSGRPSHRAVFNETHKSTRQTQVLGTQTCETLKLQPLNCIEDPVSPIVLRNF 720
Db	677	YDLALDSGRPSHRAVFNETHKSTRQTQVLGTQTCETLKLQPLNCIEDPVSPIVLRNF 736
QY	721	SLVGTPLAFNGLRPLVAEDAQRLLFTALFPPEKNGNDNICODDLSITFSEMSLDCLVVG 780
Db	737	SLVGTPLAFNGLRPLVAEDAQRLLFTALFPPEKNGNDNICODDLSITFSEMSLDCLVVG 796
QY	781	GPREENVTTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRQSRWELACESASSTEV 840
Db	797	GPREENVTTVVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNRQSRWELACESASSTEV 856
QY	841	SGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 900
Db	857	SGALKSTCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 916
QY	901	QLELPVKYAVYMWVTSYKLYNFTASENTSRVWQHGYQVSNLQSRSLPSLFLVFPV 960
Db	917	QLELPVKYAVYMWVTSYKLYNFTASENTSRVWQHGYQVSNLQSRSLPSLFLVFPV 976
QY	961	RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCSTAVCQRIQCDIP 1020

SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 5; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVANORSLYQCDYSTGSCBPI 60  
DB 17 FNLDTENAMTFOENARGFQSVVQLQGSRRVVGAPQEIIVANORSLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPPVHOTCSNTYVVKGLCFGLFSGNLRQPOK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPPVHOTCSNTYVVKGLCFGLFSGNLRQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSTIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDGSGSTIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEEF 196

QY 181 RIHFTTFEFOQNNPRSLIKPITQLLGRTHATGLRKVVRELENTINGARKNAFKILPLL 240  
DB 197 RIHFTTFEFOQNNPRSLIKPITQLLGRTHATGLRKVVRELENTINGARKNAFKILVVI 256

QY 241 TDGKFGDPLGYEDVPELDRGVIRYVGLFGDAPRSEKSKOELNTVASKPPRDHVFQAN 300  
DB 257 TDGKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSKOELNTIASKPPRDHVFQVN 316

QY 301 NFEALKTVQNLREKIFAIEGTQTGSSSSFFHEMSQEGFSAITNSGPELLSTVGSYDWAG 360  
DB 317 NFEALKTVQNLREKIFAIEGTQTGSSSSFFHEMSQEGFSAITNSGPELLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDALGYAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTVRVDSMDNDALGYAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436

QY 421 QNTGWESNANVKQTQI GAYFGASLCSDVDVDSNGSTDLVLIGAPHYEQTRGGQVSCPL 480  
DB 437 QNTGWESNANVKQTQI GAYFGASLCSDVDVDSNGSTDLVLIGAPHYEQTRGGQVSCPL 496

QY 481 PRGORARQCDVLYGQGGPWGFRGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGORARQCDVLYGQGGPWGFRGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLLRSQ 600  
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLLRSQ 616

QY 601 PVLVRKAIMBPNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
DB 617 PVLVRKAIMBPNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDVPSPVILRLNF 720  
DB 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDVPSPVILRLNF 736

QY 721 SLVGTPLUSAFGNLRPLVLAEDAQRFTALFPFEKNCNDNICQDDLSITFFSMLDCLVWG 780  
DB 737 SLVGTPLUSAFGNLRPLVLAEDAQRFTALFPFEKNCNDNICQDDLSITFFSMLDCLVWG 796

QY 781 GPREFNVTVVRNDGDSYQTVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
DB 797 GPREFNVTVVRNDGDSYQTVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGKLLKANTVSENMPRTNKTEF 900  
DB 857 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGKLLKANTVSENMPRTNKTEF 916

QY 901 QLELPVKYAVYVMVTSHGVSCTKYNFTASNTSRVMQHYQVSNLQSRSLPISLVFLVPV 960  
DB 917 QLELPVKYAVYVMVTSHGVSCTKYNFTASNTSRVMQHYQVSNLQSRSLPISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCQRIQCDIP 1020  
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVCSIAVCQRIQCDIP 1036

QY 1021 PFGIOBEFNATLKGNSPDWYIKTSHNHLILVSTAEILFNDVSFTLLPQGAFAVSQET 1080  
DB 1037 PFGIOBEFNATLKGNSPDWYIKTSHNHLILVSTAEILFNDVSFTLLPQGAFAVSQET 1096

QY 1081 KVEPEFVNPPLPLIVGSSVGGLLILALITAAALYKLGFEKROYKDMSSGGPPGAEPQ 1137  
DB 1097 KVEPEFVNPPLPLIVGSSVGGLLILALITAAALYKLGFEKROYKDMSSGGPPGAEPQ 1153

## RESULT 5

AAO14428  
ID AAO14428 standard; protein; 1153 AA.  
XX AAO14428;  
AC AAO14428;  
XX  
DT 03-MAY-2002 (first entry)  
XX Integrin Mac-1 alpha subunit.  
DE  
XX  
XX Mac-1; integrin alpha subunit; variant integrin inserted domain protein;  
KW open conformation; integrin related inflammatory disorder;  
KW integrin related immunological disorder; rheumatoid arthritis; ischaemia;  
KW reperfusion; hypovolemic shock; infarction; cerebral shock;  
KW viral infection; cancer; gene therapy; vaccine;  
KW bioactive agent screening.  
XX  
OS Unidentified.  
XX  
PN WO200204521-A2.  
XX  
PD 17-JAN-2002.  
XX  
PF 09-JUL-2001; 2001WO-US021805.  
XX  
PR 07-JUL-2000; 2000US-0216600P.  
XX  
PA (CALY ) CALIFORNIA INST OF TECHNOLOGY.  
PA (BLOO-) CENT BLOOD RES.  
XX  
PI Springer T;  
XX  
XX WPI; 2002-148167/19.  
XX  
PT New integrin I domain protein having alteration in at least 2  
PT noncontiguous regions and exits in an open conformation, useful for  
PT treating, preventing or suppressing inflammatory or immunological  
PT disorders.  
XX  
XX Example 1; Fig 1F; 90pp; English.  
PS  
XX The invention comprises structurally biased variant integrin inserted (I)  
CC domain proteins, wherein the alterations to the protein occur in at least  
CC two noncontiguous regions. Specifically the variant integrin I domain  
CC proteins are structurally biased to exist in the open conformation,  
CC thereby altering the binding ability of the protein. The invention also  
CC comprises nucleic acids encoding the variant integrin I domain proteins.  
CC The integrin I domain proteins and nucleic acids are useful for treating,  
CC preventing or suppressing integrin related inflammatory and immunological  
CC disorders (e.g. rheumatoid arthritis). The variant integrin I domain  
CC proteins and nucleic acids can also be used for treating: ischaemia/  
CC reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral  
CC infection; and cancer. The variant integrin I domain nucleic acids and  
CC proteins may be used in gene therapy, as vaccines and to screen for  
CC bioactive agents. The present amino acid sequence represents the Mac-1  
CC alpha subunit of integrin  
XX  
SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 5; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSIVOLQGSRVVVGAPQEIIVAAANORGLSYQCDYSTGSCBPI 60  
DB |||||  
QY 17 FNLDTENAMTQENARGFGQSIVOLQGSRVVVGAPQEIIVAAANORGLSYQCDYSTGSCBPI 76  
DB |||||  
QY 61 RLQVVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFNSLRQQPK 120  
DB |||||  
QY 77 RLQVVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLFNSLRQQPK 136  
DB |||||  
QY 121 FPEARLGCPEQSDIAFLIDGSGIIPHDFFRMKELVSTIMEQLKSKTSLFSLMOYSBEF 180  
DB |||||  
QY 137 FPEARLGCPEQSDIAFLIDGSGIIPHDFFRMKELVSTIMEQLKSKTSLFSLMOYSBEF 196  
DB |||||  
QY 181 RIHFTFKFQNNPNSRLKPIITOLLGRTHATGLKVVRELFNITNGARKNAFKILFL 240  
DB |||||  
QY 197 RIHFTFKFQNNPNSRLKPIITOLLGRTHATGLKVVRELFNITNGARKNAFKILVVI 256  
DB |||||  
QY 241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGFADAFRSEKSRQELNTVASKPRDHVFQAN 300  
DB |||||  
QY 257 TDGEKFGDPLGYEDVIPELDREGVIRYVIGFADAFRSEKSRQELNTVASKPRDHVFQAN 316  
DB |||||  
QY 301 NFEALKTQVONQUREKI PAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDRAG 360  
DB |||||  
QY 317 NFEALKTQVONQUREKI PAIEGTQTGSSSFEHMSQEGFSAITNSGPLLSTVGSYDRAG 376  
DB |||||  
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420  
DB |||||  
QY 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 436  
DB |||||  
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 480  
DB |||||  
QY 437 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 496  
DB |||||  
QY 481 PRGQARWQCDVAVLGEQOPNGRFGAALTVLGVNGDKLTVDAIGAPGEEDNRGAVYLF 540  
DB |||||  
QY 497 PRGQARWQCDVAVLGEQOPNGRFGAALTVLGVNGDKLTVDAIGAPGEEDNRGAVYLF 556  
DB |||||  
QY 541 HTSGSGTSPSHSQRITAGSKLPRLOYFGQSLSGGQDLTMDGLVLTGVAQGHVLLLRQ 600  
DB |||||  
QY 557 HTSGSGTSPSHSQRITAGSKLPRLOYFGQSLSGGQDLTMDGLVLTGVAQGHVLLLRQ 616  
DB |||||  
QY 601 PVLRVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRLEGQIQSVVT 660  
DB |||||  
QY 617 PVLRVKAIMFENPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLRLEGQIQSVVT 676  
DB |||||  
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGTQTCETIKLQPLNCIEDPVPVILRLNF 720  
DB |||||  
QY 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGTQTCETIKLQPLNCIEDPVPVILRLNF 736  
DB |||||  
QY 721 SLVGTPLSAFGLNLRPVLAEDAQLFTALPFPEKNCNDNI CODDLSTITFSFMSLCLVVG 780  
DB |||||  
QY 737 SLVGTPLSAFGLNLRPVLAEDAQLFTALPFPEKNCNDNI CODDLSTITFSFMSLCLVVG 796  
DB |||||  
QY 781 GPREFNVTVVRNDSYRTQVTPFPFLDLVSRKYSTLQNSQRSWRLACESASSTEV 840  
DB |||||  
QY 797 GPREFNVTVVRNDSYRTQVTPFPFLDLVSRKYSTLQNSQRSWRLACESASSTEV 856  
DB |||||  
QY 841 SGALKTSKCSINHPIPENSEVFNITFDVDSKASLGNKLLKANTVSENPNRTNKTTF 900  
DB |||||  
QY 857 SGALKTSKCSINHPIPENSEVFNITFDVDSKASLGNKLLKANTVSENPNRTNKTTF 916  
DB |||||  
QY 901 QLELPVKYAVYVMTVSHGVSTKYNFTASENTSRVQHOYQVSNLQGRSLPISLVLFPV 960  
DB |||||  
QY 917 QLELPVKYAVYVMTVSHGVSTKYNFTASENTSRVQHOYQVSNLQGRSLPISLVLFPV 976  
DB |||||  
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSPSHSDFLAELRKAPVNCSTIACVQRIQCDIP 1020  
DB |||||  
QY 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSPSHSDFLAELRKAPVNCSTIACVQRIQCDIP 1036  
DB |||||  
QY 1021 FFGIOBEFNATLKNLSFDWYIKTSNNHLLIVSTAEILLFNDVSTFTLLPGQAFVRSQTET 1080  
DB |||||  
QY 1037 FFGIOBEFNATLKNLSFDWYIKTSNNHLLIVSTAEILLFNDVSTFTLLPGQAFVRSQTET 1096  
DB |||||  
QY 1081 KVEPFFVEPNPLPLVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPFGAEPQ 1137

DB 1097 KVEPFFVEPNPLPLVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPFGAEPQ 1153  
|||  
RESULT 6  
ADD25615  
ID ADD25615 standard; protein; 1153 AA.  
XX  
AC ADD25615;  
XX  
DT 15-JAN-2004 (first entry)  
XX  
DE Binding domain-immunoglobulin fusion protein-associated protein #85.  
XX  
DE Binding domain; immunoglobulin; fusion protein; cytostatic;  
KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
KW neuroprotective; hinge region; immunoglobulin heavy chain;  
KW CH2 constant region; CH3 constant region; IgG1;  
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.  
XX  
OS Unidentified.  
XX  
XX US2003118592-A1.  
PN  
XX  
PD 26-JUN-2003.  
XX  
XX 25-JUL-2002; 2002US-00207655.  
PF  
XX  
PR 17-JAN-2001; 2001US-0367350P.  
PR 17-JAN-2002; 2002US-00053530.  
PR 03-JUN-2002; 2002US-0385691P.  
XX  
PA (GENE-) GENE-CRAFT INC.  
XX  
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
PI WPI; 2003-801317/75.  
DR  
XX  
XX New binding domain-immunoglobulin fusion protein, useful for treating a  
PT subject having or suspected of having a malignant condition or a B-cell  
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
XX  
PS Disclosure; SEQ ID NO 176; 157pp; English.  
XX  
CC The invention relates to a binding domain-immunoglobulin fusion protein  
CC comprising a binding domain polypeptide that is fused to an  
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
CC CH2 constant region polypeptide that is fused to the hinge region  
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region  
CC polypeptide that is fused to the CH2 constant region polypeptide. The  
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
CC region polypeptide, derived from (a) having 3 or more cysteine residues;  
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
CC immunoglobulin hinge region polypeptide contains no more than one  
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
CC polypeptide, derived from (a) having 3 or more cysteine residues; where  
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
CC capable of at least one immunological activity comprising antibody  
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
CC binding domain polypeptide is capable of specifically binding to an  
CC antigen. Also included are an isolated polynucleotide encoding the  
CC binding domain-immunoglobulin fusion protein, a recombinant expression  
CC construct comprising the polynucleotide (operably linked to a promoter),  
CC a host cell transformed or transfected with a recombinant expression  
CC construct, producing the binding domain-immunoglobulin fusion protein, a

CC pharmaceutical composition comprising the binding domain-immunoglobulin  
 CC fusion protein or polynucleotide and a carrier, and treating a subject  
 CC having or suspected of having a malignant condition or a B-cell disorder.  
 CC The binding domain-immunoglobulin fusion protein is useful for treating a  
 CC subject having or suspected of having a malignant condition or a B-cell  
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple  
 CC sclerosis or autoimmune disease. The present sequence is a binding domain  
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
 CC sequence data for this patent formed part of the printed specification  
 CC and is also available in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not  
 CC identified the sequences in the printed specification by their SEQ ID  
 CC number therefore none of the sequences can be explicitly identified.

XX  
 SQ Sequence 1153 AA;

Query Match 99.3%; Score 5839; DB 7; Length 1153;  
 Best Local Similarity 98.9%; Pred. No. 0;  
 Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVLOGSRVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 60  
 DB 17 FNLDTENAMTFQENARGFGQSVVLOGSRVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 76  
 QY 61 RLQVPVEAVNMVLSGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOQK 120  
 DB 77 RLQVPVEAVNMVLSGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOQK 136  
 QY 121 FPEALRGCPQEDSDIAFLDGSIIIPDPRMKELVSTIMEQLKSKTFLSLMQYSEEF 180  
 DB 137 FPEALRGCPQEDSDIAFLDGSIIIPDPRMKEFVSTVMEQLKSKTFLSLMQYSEEF 196  
 QY 181 RIHFTKEFQNNPRSLVPIPTQLLGRTHATGLRKVRELFNITNGARKNAPKILFL 240  
 DB 197 RIHFTKEFQNNPRSLVPIPTQLLGRTHATGLRKVRELFNITNGARKNAPKILVWI 256  
 QY 241 TDEKFGDPLGYEDVIEPDLREGVIRVVLGFGDAFRSEKSRQELNTVASKEPRDHVFQAN 300  
 DB 257 TDEKFGDPLGYEDVIEPDLREGVIRVVLGFGDAFRSEKSRQELNTIASKEPRDHVFQVN 316  
 QY 301 NFPAKTVQNLREKIFAIEGTQTGSSSFHEHMSQEGFSAATNSGPFLLSTVGSYDWAG 360  
 DB 317 NFPAKTVQNLREKIFAIEGTQTGSSSFHEHMSQEGFSAATNSGPFLLSTVGSYDWAG 376  
 QY 361 GVFLYTSKEKSTINMTRVDSMDNDAYLGAAAILLRNVOSIVLGAPRYOHIGLVAMFR 420  
 DB 377 GVFLYTSKEKSTINMTRVDSMDNDAYLGAAAILLRNVQSLVIGAPRYOHIGLVAMFR 436  
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSCPL 480  
 DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSCPL 496  
 QY 481 PRQORARWQCDVLYGEGQPGWRFGAALTDLVGDVNGDKLTDVAIGAPGEEDNRGAYL 540  
 DB 497 PRQORARWQCDVLYGEGQPGWRFGAALTDLVGDVNGDKLTDVAIGAPGEEDNRGAYL 556  
 QY 541 HGTSGSGISPHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
 DB 557 HGTSGSGISPHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
 QY 601 PVLVRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTEDRLREGIOISVVT 660  
 DB 617 PVLVRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTEDRLREGIOISVVT 676  
 QY 661 YDLALDSGRPHSAVFNFTKNSTRQTQVLGLTQTCETLKLQPNCLIEDPVSPIVLRNLF 720  
 DB 677 YDLALDSGRPHSAVFNFTKNSTRQTQVLGLTQTCETLKLQPNCLIEDPVSPIVLRNLF 736  
 QY 721 SLVGTPLSAFGNLRPVLAEDAQRILFTALFPPEKNCNDNICQDDLSTITFSFMSLDCI 780  
 DB 737 SLVGTPLSAFGNLRPVLAEDAQRILFTALFPPEKNCNDNICQDDLSTITFSFMSLDCI 796

QY 781 GPREENVTIVRNDGEDSYRTQVTFPPDLVSRKSVSTLQNSORSORSWRLACESASSTEV 840  
 DB 797 GPREENVTIVRNDGEDSYRTQVTFPPDLVSRKSVSTLQNSORSORSWRLACESASSTEV 856  
 QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900  
 DB 857 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916  
 QY 901 QLELPVKYAYVMVTSHGVSSTKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPV 960  
 DB 917 QLELPVKYAYVMVTSHGVSSTKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPV 976  
 QY 961 RLNQTVLWDRPQVTFSENLSSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQC 1020  
 DB 977 RLNQTVLWDRPQVTFSENLSSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQC 1036  
 QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTET 1080  
 DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSOTET 1096  
 QY 1081 KVEPFEVNPDLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKXDMMEGPPGAPQ 1137  
 DB 1097 KVEPFEVNPDLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKXDMMEGPPGAPQ 1153  
 RESULT 7  
 AAR04136  
 ID AAR04136 standard; protein; 1153 AA.  
 XX AAR04136;  
 AC AC  
 DT 09-SEP-2004 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 07-SEP-1990 (first entry)  
 XX  
 DE Alpha subunit of Mac-1 leukocyte adhesion receptor.  
 XX  
 KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;  
 KW non-specific defence system; integrin gene superfamily.  
 XX  
 OS Synthetic.  
 FH  
 FT Location/Qualifiers  
 FT Region 1..116  
 FT /label= signal\_peptide  
 FT Modified-site 86..88  
 FT /label= putative N-glycosylation site  
 FT Modified-site 240..242  
 FT /label= putative N-glycosylation site  
 FT Modified-site 391..393  
 FT /label= putative N-glycosylation site  
 FT Modified-site 469..471  
 FT /label= putative N-glycosylation site  
 FT Modified-site 693..695  
 FT /label= putative N-glycosylation site  
 FT Modified-site 697..699  
 FT /label= putative N-glycosylation site  
 FT Modified-site 735..737  
 FT /label= putative N-glycosylation site  
 FT Modified-site 802..804  
 FT /label= putative N-glycosylation site  
 FT Modified-site 881..883  
 FT /label= putative N-glycosylation site  
 FT Modified-site 901..903  
 FT /label= putative N-glycosylation site  
 FT Modified-site 912..914  
 FT /label= putative N-glycosylation site  
 FT Modified-site 941..943  
 FT /label= putative N-glycosylation site  
 FT Modified-site 947..949  
 FT /label= putative N-glycosylation site  
 FT Modified-site 979..981  
 FT /label= putative N-glycosylation site

FT Modified-site 994..996  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1022..1024  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1045..1047  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1051..1053  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1076..1078  
 FT /label= putative N-glycosylation site  
 FT Region 1106..1134  
 FT /label= putative\_transmembrane\_region  
 XX  
 PN EP364690-A.  
 XX  
 PD 25-APR-1990.  
 XX  
 PF 17-AUG-1989; 89EP-00115159.  
 XX  
 XX 23-AUG-1988; 88US-00235353.  
 PR 09-MAR-1989; 89US-00321239.  
 PR  
 XX (DAND ) DANA FARBER CANCER INST INC.  
 PA  
 XX Springer TA, Corbi A;  
 PI  
 XX WPI: 1990-125938/17.  
 DR N-PSDB; AAQ04043.  
 DR  
 XX New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating  
 PT inflammation and viral infections, and in diagnosis.  
 PT  
 XX Disclosure; Page ?; 3pp; English.  
 PS  
 XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.  
 CC recognition of and migration to sites of inflammation. It also attaches  
 CC to cellular substrates as part of this function making it useful in  
 CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene  
 CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25  
 CC -MAR-2003 to correct PA field.)  
 CC  
 CC Revised record issued on 09-SEP-2004 : Correction to feature table key  
 CC and pages  
 CC  
 XX Sequence 1153 AA;  
 SQ  
 Query Match 99.1%; Score 5829; DB 2; Length 1153;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 1123; Conservative 8; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 ENLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGSLYQCDYSTGSCBPI 60  
 DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAANQSGSLYQCDYSTGSCBPI 76  
 QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFGSLNRQQPQK 120  
 DB 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFLFGSLNRQQPQK 136  
 QY 121 PPEALRGCPQESDIAFLIDGSGSIIPDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 180  
 DB 137 PPEALRGCPQESDIAFLIDGSGSIIPDPRMKELVSTIMEQLKSKTFLSLMOYSEEF 196  
 QY 181 RHFTFKFQNNPNSRLKPTQLLGRTHTTATGKRVRELFNITNGARKNAKILFL 240  
 DB 197 RHFTFKFQNNPNSRLKPTQLLGRTHTTATGKRVRELFNITNGARKNAKILVVI 256  
 QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDRHVQAN 300  
 DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDRHVQAN 316  
 QY 301 NFEALKTQVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNGPLLSTVGYDWDAG 360  
 DB 317 NFEALKTQVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNGPLLSTVGYDWDAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVOSLVLGAPRYOHIGLVAMFR 420  
 DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVOSLVLGAPRYOHIGLVAMFR 436  
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLCAPHYEQTRGGQVSVCP 480  
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLCAPHYEQTRGGQVSVCP 496  
 QY 481 PRGQARWQCDVILYGEQGPWGRFGAALTVLGVGVGDKLTDVAIGAPGEDNAGVYLF 540  
 DB 497 PRGQARWQCDVILYGEQGPWGRFGAALTVLGVGVGDKLTDVAIGAPGEDNAGVYLF 556  
 QY 541 HGTSGSISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVNDLTVGAQCHVLLLSRQ 600  
 DB 557 HGTSGSISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVNDLTVGAQCHVLLLSRQ 616  
 QY 601 PVLRVKAIMFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRRLREGQIQSVVT 660  
 DB 617 PVLRVKAIMFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRRLREGQIQSVVT 676  
 QY 661 YDLALDSGRPHSRVAFNETKNSRTOVGLGTOTCETLKLQLPNCIEDPVPVIRLNF 720  
 DB 677 YDLALDSGRPHSRVAFNETKNSRTOVGLGTOTCETLKLQLPNCIEDPVPVIRLNF 736  
 QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNI QDDLSITFSFMSLDCLVVG 780  
 DB 737 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNI QDDLSITFSFMSLDCLVVG 796  
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
 DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856  
 QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
 DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916  
 QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHQYQVSNLQSRSLPISLVFLVVP 960  
 DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHQYQVSNLQSRSLPISLVFLVVP 976  
 QY 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKPAPVNVVNCISVAVCQRIQCDIP 1020  
 DB 977 RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKPAPVNVVNCISVAVCQRIQCDIP 1036  
 QY 1021 PFGIQEEFNATLKGNLSEFDWIKTSHNHLIVSTAEILFNDSVFTLLPGQGFVRSQDET 1080  
 DB 1037 PFGIQEEFNATLKGNLSEFDWIKTSHNHLIVSTAEILFNDSVFTLLPGQGFVRSQDET 1096  
 QY 1081 KVEPPEVENPLPLIVGSSVGGLLILALITAAALYKLGFFKROYKDMMSGGPPGABPPQ 1137  
 DB 1097 KVEPPEVENPLPLIVGSSVGGLLILALITAAALYKLGFFKROYKDMMSGGPPGABPPQ 1153  
 RESULT 8  
 ADM99589  
 ID ADM99589 standard; protein; 1152 AA.  
 XX  
 AC ADM99589;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Human integrin alphaM subunit precursor protein.  
 XX  
 KW integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;  
 KW osteopathic; cytostatic; immunosuppressive; antiinflammatory;  
 KW neuroprotective; antiskinning; immunotherapy; inflammatory;  
 KW autoimmune disorder; thrombosis; cancer; osteoporosis;  
 KW sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;  
 KW alphaM.  
 XX  
 OS Homo sapiens.  
 XX

PH	Key	Location/Qualifiers	
FT	Misc-difference 965	/note= "Encoded by CCC"	
FT			
XX			
PN	WO2004007530-A2.		
XX			
PD	22-JAN-2004.		
XX			
XX	17-JUL-2003; 2003WO-US022301.		
PF			
XX			
PR	17-JUL-2002; 2002US-0396783P.		
PR	17-JUL-2002; 2002US-0396790P.		
PR	11-SEP-2002; 2002US-0410135P.		
XX			
XX	(BLOO-) CENT BLOOD RES INC.		
PA			
XX			
PI	Springer TA, Takagi J;		
XX			
DR	WPI; 2004-122877/12.		
DR	N-PSDB; ADM99588.		
XX			
PT	Novel modified integrin protein having extracellular domains of integrin alpha and beta subunits or integrin alphaII and beta3 subunit, useful for treating integrin mediated disorders.		
PT			
PT			
XX			
PS	Disclosure; SEQ ID NO 4; 232pp; English.		
XX			
CC	The invention relates to a novel isolated or recombinant modified integrin protein having extracellular domains of integrin alpha and beta subunits where one of the subunits has one or more mutations, an altered surface feature or an amino acid substitution or internal deletion, extracellular domains of the integrin beta subunit that comprise a mutation that alters a non-cysteine residue to cysteine or extracellular domains of integrin alpha and beta subunits. The polypeptide of the invention demonstrates antiproliferative, thrombolytic, anticoagulant, osteoprotective, cyostatic, immunosuppressive, antiinflammatory, neuroprotective and antiskinning activities and may be useful for immunotherapy in order to prevent or treat an integrin-mediated disorder such as an inflammatory disorder, an autoimmune disorder, thrombosis, cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple sclerosis. The current sequence is that of the human integrin alphaM subunit precursor protein of the invention.		
XX			
XX	Sequence 1152 AA;		
	Query Match	99.1%; Score 5823.5; DB 8; Length 1152;	
	Best local Similarity	98.8%; Pred. No. 0;	
	Matches 1123; Conservative	8; Mismatches	5; Indels
			1; Gaps
QY	1	FNLDTENAMTFOENARGFGQSVVQLOGSRVVGAPQEIIVANQSGSYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFOENARGFGQSVVQLOGSRVVGAPQEIIVANQSGSYQCDYSTGSCPEI	76
QY	61	RLQVPVEAVNMISGLSLAATTPSQLLACGPTVHQTCSENTYVKGLCFGLGSLNRQOPQK	120
DB	77	RLQVPVEAVNMISGLSLAATTPSQLLACGPTVHQTCSENTYVKGLCFGLGSLNRQOPQK	136
QY	121	FPEALRCQPEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTFLPSLMQYSEEF	180
DB	137	FPEALRCQPEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTFLPSLMQYSEEF	196
QY	181	RIHFTFKFQNNPRLSKIPITOLLGRTHATGLRKVRELFINITNGARKNAFKILFL	240
DB	197	RIHFTFKFQNNPRLSKIPITOLLGRTHATGLRKVRELFINITNGARKNAFKILVVI	256
QY	241	TDGEKFGDPLGYEDVPELDRGVIRYVVGDAFRSEKSRQELNTVASKPPRDHVFQAN	300
DB	257	TDGEKFGDPLGYEDVPEADREGVIRYVVGDAFRSEKSRQELNTVASKPPRDHVFQVN	316
QY	301	NFEALKTVQNLREKIPFAIGTQTGSSSSFEHEMSQEGFSAATNSGPLLSTVGSYDWAG	360
DB	317	NFEALKTVQNLREKIPFAIGTQTGSSSSFEHEMSQEGFSAATNSGPLLSTVGSYDWAG	376
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMFR	420
DB	377	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYOHIGLVAMFR	436
QY	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL	480
DB	437	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSCPL	496
QY	481	PRGORARWQCDAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF	540
DB	497	PRG-RARWQCDAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVYLF	555
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGDLTMDGLVDLTVGAGHVLRLRSQ	600
DB	556	HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSGGDLTMDGLVDLTVGAGHVLRLRSQ	615
QY	601	PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRLVHVKSTRDLREGIOQSVVT	660
DB	616	PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRLVHVKSTRDLREGIOQSVVT	675
QY	661	YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRINF	720
DB	676	YDLALDSGRPHSRVAFNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRINF	735
QY	721	SLVGTPLSAFGLRPLVLAEDAQRLEFALFPFEXKCGNDNICODDLSTIFFSFMSLDCLVVG	780
DB	736	SLVGTPLSAFGLRPLVLAEDAQRLEFALFPFEXKCGNDNICODDLSTIFFSFMSLDCLVVG	795
QY	781	GPREFNVTVVRNDGEDSYRTQVTFPPPLDLVYRKVSTLQNRQSRWRLACESASSTEV	840
DB	796	GPREFNVTVVRNDGEDSYRTQVTFPPPLDLVYRKVSTLQNRQSRWRLACESASSTEV	855
QY	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
DB	856	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	915
QY	901	QLELPVKYAVYVMTSHGVSTKVLNFTASENTSRVMQHOYQVSNLGRSLPISLVLVLPV	960
DB	916	QLELPVKYAVYVMTSHGVSTKVLNFTASENTSRVMQHOYQVSNLGRSLPISLVLVLPV	975
QY	961	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCORIQCDIP	1020
DB	976	RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCORIQCDIP	1035
QY	1021	FFGIQEEFNATLKNLSFDWYIKTSNNHLLIVSTABILFNDSVPTLLPGQAFVRSQTET	1080
DB	1036	FFGIQEEFNATLKNLSFDWYIKTSNNHLLIVSTABILFNDSVPTLLPGQAFVRSQTET	1095
QY	1081	KVEPFEVNPPLIVGSSVGLLALLALITAAALYKLGFFKQYKDMKSEGGPPGAEPQ	1137
DB	1096	KVEPFEVNPPLIVGSSVGLLALLALITAAALYKLGFFKQYKDMKSEGGPPGAEPQ	1152
	RESULT 9		
	ADP12435		
ID	ADP12435	standard; protein; 1152 AA.	
XX	ADP12435;		
AC	ADP12435;		
XX			
DT	12-AUG-2004 (first entry)		
XX			
DE	Protein encoded by mRNA of the invention #45.		
XX			
KW	transplant rejection; immune system; rheumatoid arthritis; lupus;		
XX	inflammatory bowel disease; multiple sclerosis; HIV; AIDS.		
OS	Homo sapiens.		
XX			
PN	WO2004042346-A2.		
PD	21-MAY-2004.		
XX			
PF	24-APR-2003; 2003WO-US012946.		

XX 24-APR-2002; 2002US-00131831.  
PR 20-DEC-2002; 2002US-00325899.  
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.  
PA Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;  
PI Rosenberg S;  
XX WPI; 2004-400724/37.  
DR Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
XX pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
PT rejection, in an individual, comprises detecting the expression level of  
PT the genes.  
XX  
PS Claim 65; SEQ ID NO 2444; 1762pp; English.  
XX The present invention relates to diagnosing or monitoring transplant  
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual  
CC comprises detecting the expression level of one or more genes. The  
CC methods, system and kits are useful in diagnosing or monitoring  
CC islet, lung, bone marrow or stem cell transplant rejection,  
CC xenotransplant rejection or mechanical organ replacement rejection, in an  
CC individual. The methods are also useful in assessing the immune status of  
CC an individual. The methods are also useful in diagnosing and monitoring  
CC diseases that involve the immune system, e.g. rheumatoid arthritis,  
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
CC viral, bacterial or fungal infection. The present sequence represents a  
CC protein that is encoded by the mRNA of the invention.  
XX  
SQ Sequence 1152 AA;  
Query Match 98.9%; Score 5816.5; DB 8; Length 1152;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1122; Conservative 8; Mismatches 6; Indels 1; Gaps 1;  
QY 1 ENLDTENAMTFOENARGGQSVVLOGGVRVVGAPQEIIVAAANQGSLSYQCDYSTGSCPEI 60  
DB 17 ENLDTENAMTFOENARGGQSVVLOGGVRVVGAPQEIIVAAANQGSLSYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKGLCFGLGSLNRQQPQK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKGLCFGLGSLNRQQPQK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKXKTLPSLMQYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKXKTLPSLMQYSEEF 196  
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGKRVRELFTNITNGARKNAFKILFL 240  
DB 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGKRVRELFTNITNGARKNAFKILVVI 256  
QY 241 TDGEFGPGLGYEDVPELDRGVIRYVLGFGDAFRSEKSRQELNNTVASKPPRDRHVQAN 300  
DB 257 TDGEFGPGLGYEDVPEADRGVIRYVIGVGDAPFRSEKSRQELNNTVASKPPRDRHVQAN 316  
QY 301 NPEALKTVQNLQREKIFAIEGTQTGSSSSFEHMSQEGHSAITNSGPLLSVTGSDWAG 360  
DB 317 NPEALKTVQNLQREKIFAIEGTQTGSSSSFEHMSQEGHSAITNSGPLLSVTGSDWAG 376  
QY 361 GVFLYTSKESKTFINNRVDSMDNDAYLGAAAILRNVRQSLVLGAPYQHIGLVAMPR 420  
DB 377 GVFLYTSKESKTFINNRVDSMDNDAYLGAAAILRNVRQSLVLGAPYQHIGLVAMPR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYYEQTRGGQSVCP 480  
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYYEQTRGGQSVCP 496  
QY 481 PRGQARWQCDVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVILF 540  
DB 497 PRG-RARWQCDVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVILF 555

QY 541 HGTSGSGTSPSHSRIAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTVGAQHVLILRSQ 600  
DB 556 HGTSGSGTSPSHSRIAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTVGAQHVLILRSQ 615  
QY 601 FVLRVKAIIMEFNPREVARNFECNDQVVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 616 FVLRVKAIIMEFNPREVARNFECNDQVVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675  
QY 661 YDLALDSGRPHSRVAFNETKSTRTOVLGTLTCTCETLKLQLPNCIEDPSPVILRNF 720  
DB 676 YDLALDSGRPHSRVAFNETKSTRTOVLGTLTCTCETLKLQLPNCIEDPSPVILRNF 735  
QY 721 SLVGTPLSAFGNLRPLVLAEDAQRULFTALFPPEKKNCGNDNICDDLSITFSFMSLDCLVVG 780  
DB 736 SLVGTPLSAFGNLRPLVLAEDAQRULFTALFPPEKKNCGNDNICDDLSITFSFMSLDCLVVG 795  
QY 781 GPRFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840  
DB 796 GPRFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 855  
QY 841 SGALKSTSCSINHPIFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
DB 856 SGALKSTSCSINHPIFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915  
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQSRSLPISLVFLVPV 960  
DB 916 QLELPVKYAVVMVTVSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQSRSPISLVFLVPV 975  
QY 961 RLNQTVIWDREPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCRIQCIP 1020  
DB 976 RLNQTVIWDREPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCRIQCIP 1035  
QY 1021 PFGIQEEENATLKGNLSEFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQTE 1080  
DB 1036 PFGIQEEENATLKGNLSEFDWIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSQTE 1095  
QY 1081 KVEPPEVENPPLIIVGSSVGGLLIALITALIYKLGFFKQYKDMWSEGGPPGABPQ 1137  
DB 1096 KVEPPEVENPPLIIVGSSVGGLLIALITALIYKLGFFKQYKDMWSEGGPPGABPQ 1152  
RESULT 10  
ADP44061  
ID ADP44061 standard; protein; 1163 AA.  
XX ADP44061;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
XX Human CD11C protein SEQ ID NO:14.  
DE  
XX  
KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;  
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.  
XX  
XX Homo sapiens.  
XX  
XX WO2004053094-A2.  
PN  
XX  
PD 24-JUN-2004.  
XX  
PF 08-DEC-2003; 2003WO-US039208.  
XX  
PR 06-DEC-2002; 2002US-0431522P.  
XX  
PA (PPDP-) PPD DEV LP.  
XX  
XX Dunn SJ, Holzwayer TA;  
PI  
XX  
XX WPI: 2004-480928/45.  
DR N-PSDB; ADP44060.  
XX  
XX Identifying an inhibitor of human immunodeficiency virus (HIV) entry into



a human host cell useful for preventing and/or treating HIV infection, by identifying an inhibitor of a cell surface polypeptide such as CXCR-4.

Claim 1; SEQ ID NO 14; 133pp; English.

The present invention describes a method (M1) for identifying an inhibitor of HIV entry into a human host cell. (M1) comprises identifying an inhibitor of a cell surface polypeptide selected from CXCR-4 (352 amino acids, SEQ ID NO:8, ADP44055), CCR4 (360 amino acids, SEQ ID NO:10, ADP44057), CCR7 (378 amino acids, SEQ ID NO:12, ADP44059), CD11c (1163 amino acids, SEQ ID NO:14, ADP44061), CD47 (323 amino acids, SEQ ID NO:16, ADP44063), CD68 (354 amino acids, SEQ ID NO:18, ADP44065), CD69 (199 amino acids, SEQ ID NO:20, ADP44067), CD74 (566 amino acids, SEQ ID NO:22, ADP44069), CSF3R (836 amino acids, SEQ ID NO:24, ADP44071), RARA (462 amino acids, SEQ ID NO:26, ADP44073), GABBR1 (578 amino acids, SEQ ID NO:28, ADP44075), P2X1 (150 amino acids, SEQ ID NO:30, ADP44077), HELO1 (299 amino acids, SEQ ID NO:32, ADP44079), GPRK6 (576 amino acids, SEQ ID NO:34, ADP44081), or PTK2B (1009 amino acids, SEQ ID NO:36, ADP44083). Also described: (1) an inhibitor of a cell surface polypeptide as described above in a human host cell preventing HIV entry into the human host cell; (2) a pharmaceutical composition comprising an inhibitor of (1) and a carrier; and (3) conferring resistance to HIV infection in an individual, by administering the pharmaceutical composition of (2). An inhibitor of HIV entry has anti-HIV and virucide activities, and can be used as an HIV uptake inhibitor. (M1) is useful for identifying protective compounds that inhibit entry of HIV into cells, useful for the prevention and/or treatment of HIV infection. The present sequence represents human integrin alpha X (CD11c), which is used in the exemplification of the present invention.

XX Sequence 1163 AA;

Query Match 58.9%; Score 3460; DB 8; Length 1163;

Best Local Similarity 60.74; Pred. No. 3.7e-277;

Matches 685; Conservative 143; Mismatches 295; Indels 6; Gaps 4;

1 FNLDTENAMTFOBNARGFQSVQLOGSRVVGAPQEIIVAAORGSLDYQDSTGCEPI 60  
 20 FNLDTTELTAFRVDSAGFSDVQYANVWVGAPQKITAAQTGGLYQCGYSTGCEPI 79  
 61 RLQVPEAVNMSLGLSLAATSPQLLACGPTVHTQCSNTYVKGCLFLGSLNRQOPQK 120  
 80 GLQVPEAVNMSLGLSLAATSPQLLACGPTVHTQCSNTYVKGCLFLGSLNRQOPQK 137  
 121 FPEALRGCPEDSDIAFLDGSGLIPDPRMKELVSTIMEQKSKTLFSLMQYSEFP 180  
 138 LPVSRQECPEQODIVFLDGSGLISGRNFATMNFVRAVISQFQRPSTQFSLMQFSNK 197  
 181 RIHFTKEFQNNPNRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240  
 198 QTHFTFEFRSSNPLSLASVHLQCGFTTATAIQNVRELPHASVYGARDAAKILIVI 257  
 241 TDGEKFGDPLGYEDVPELDEGVIRYVLFGDAPRSEKSRQBLNTVASKPPRDHVFQAN 300  
 258 TDGKKGDSLDYKDVIPMDAAGIIRYAIGVLAQNRNSWKLNDIASQPSQEHIFKVE 317  
 301 NFPAKTVQNLREKIFAIEGTQSGSSSEHEMSQEGFSAATISNGPLLSSTVGSVDWAG 360  
 318 DFDALDIQNLREKIFAIEGTQSGSSSEHEMSQEGFSAATISNGPLLSSTVGSVDWAG 377  
 361 GVFLYTSKEKSTINTRVDSMDNDAVLGYAAAIILNRVQSLVGLAPRYOHILGLVAMFR 420  
 378 GAFLYPNMSPPTINNSQENVMDRDSYLGSTELALWKGVQSLVGLAPRYOHTGKAVIFT 437  
 421 QNTGMMESNANVKTQIGAFYGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQSVYCP 480  
 438 QVSRQWRMKAETGTQIGSYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQSVYCP 497  
 481 PRQGRARWQCDVLYGQGPQWRFGAALTVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540  
 498 PRQWR-RWWCDVLYGQGPQWRFGAALTVLGVNGDKLTDVVGAPGGEENRGAVYLF 556  
 541 HGTSGSGISPSHSQRTAGSKSLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600

Db 557 HGVGPSISPSHSQRTAGSKSLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
 QY 601 PVLVRKAIMEFNPREVARNVFECDNQVKGKEAGEVRVCLHVQKSTRDRREGIQSVWT 660  
 Db 617 PVLWGVSMQFIPAEIPRSAFEQVVSQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676  
 QY 661 YDLALDSGRPHSAVNFYKNSRTRQVGLTQTQETLKLQPNCTIEDPSPVILRLNF 720  
 Db 677 LDIALDPGLSPRATQETKNRSLRVRLGLKAHCENFNLLPSCVEDSVTPTTLRLNF 736  
 QY 721 SLVGTPLSAFNGRLPVLAEDAQRLLFTALPFPKNCNDNICQDLSITTFSPMSLDCLVVG 780  
 Db 737 TLVKGPELLAFNRUPMLAADAQRYFTASLDPFKNCCGADHICQDNLGSLFSFPGLSLVG 796  
 QY 781 GPREFNVTVTRNDGEDSVYRTQVTFPPFLDLSYKRVSTLQNSQRSWRLACSSASSTEV 840  
 Db 797 SNLELNAEVMVMWMDGEDSVYRTQVTFPPFLDLSYKRVSTLQNSQRSWRLACSSASSTEV 854  
 QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKXANTVSENMPRTNKTEF 900  
 Db 855 SQGTWSTSCRINHLPFGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPTSKTTF 914  
 QY 901 QLELPKYAVYVMTSHGVSTKYLNTAS-ENTSRVMOHOYQVSNLQORSLPISLVFLVP 959  
 Db 915 QLELPKYAVYVMTSHGVSTKYLNTAS-ENTSRVMOHOYQVSNLQORSLPISLVFLVP 974  
 QY 960 VRLNQTVIWDPRQVTFSENLSSTCHTKERLPSHDSFLAELRKAPVNVCSIAVQRIQCDI 1019  
 Db 975 VELNQAQVMDVEVSHPNQPSLRCSEKTIAPPASDFLAHIQKNPVLDCSLAGCLRPRCDV 1034  
 QY 1020 PFGIQEEFNATLKNLSFDWIKTSHNHLILVSTAEILLPNDVSVFILLPQCGAFVRSOTE 1079  
 Db 1035 PSFSVQEEFLDFTLKNLSFGWVRQILQKXSVSVVAEITFDTSVYSQLPQGEAFMAQTT 1094  
 QY 1080 TKVEPEVENPLPLIVSGSVGGELLALITAAALYKLGFFKRYKDMWSE 1128  
 Db 1095 TVLEKYKHVNPPLIVSGSVGGELLALITAAALYKLGFFKRYKDMWSE 1143

RESULT 11  
 ADNC02004  
 ID ADNC02004 standard; protein; 1163 AA.  
 AC ADNC02004;  
 DT 17-JUN-2004 (first entry)  
 XX Human inflammatory bowel disease marker - CD11c protein.  
 DE marker gene; inflammatory bowel disease; IBD; FcgammaR IIa;  
 KW FcgammaR IIb; Mig; NRG-2; hexokinase; HM74; REG III; LPAP; Mip-1(beta);  
 KW L-selectin; EGFL6; IDO; IL-8; CD11c; TLR2; ulcerative colitis;  
 KW Crohn's disease.  
 OS Homo sapiens.  
 XX JF2004065120-A.  
 PN 04-MAR-2004.  
 XX 07-AUG-2002; 2002JP-00229705.  
 PF 07-AUG-2002; 2002JP-00229705.  
 PR (SUMI) SUMITOMO SEIYAKU KK.  
 XX WPI; 2004-209124/20.  
 DR N-PSDB; ADNC01990.  
 XX Novel marker gene of inflammatory bowel disease (IBD) comprising base  
 PT sequence of e.g., Fc gamma R IIa, Fc gamma R IIb, Mig, NRG-2,  
 hexokinase 3, HM74, CD11c, TLR2 gene, useful for screening therapeutic



Query Match		58.6%; Score 3446; DB 8; Length 1163;
Best Local Similarity		60.5%; Pred. No. 5, 4e-276;
Matches		683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;
QY	1	FNLTENAMTQENARGGQSVVQLOGSRVVVVGAPQEIIVAAQNRGLYQCQDYSTGSCPEI 60
Db	20	FNLTTEELTAFRVDSAGSDSVVQVANGSVVVGAPKITAANQTGGLYQCYGSGACEPI 79
QY	61	RLOVPVAVNMNLGSLAATTSPOLLACGTVHQTCSNTYVVGKLCFLFGSNLRQOPQK 120
Db	80	GLQVPPEAVNMNLGSLASTTSPOLLACGTVHHECGRNMYLTGLCFLGPT--QLTQR 137
QY	121	FPEALRGCPQDSDIATLIDSGSIIIPHDFFRMKELVSTIMEQLKSKTFLSLMOYSEEF 180
Db	138	LPVSRQECPRQEQDIVLIDSGSISSENFMNFRVRAVISQFQRPSTQSLMQFSNKF 197
QY	181	RIHTTFKEFQNNPNRSLIKPITOLLGTHTATGLRKVKVRELFMNITNGARKNAFKILFL 240
Db	198	QTHFTFEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257
QY	241	TGGEKFGDPLGYDIVPELDREGVIRYVVGFDARFSEKSRQELNIVASKPRDRHVQAN 300
Db	258	TGKKKGDSLDYKDVIMADAGIIRVAIGVLGAPQNRNSWKELNDTASKEPSQBIKFVE 317
QY	301	NFEAKTVQNLREKIFAIEGTQTCSSSSSFHEMSQEGFSAATISNGPLLTGVSXYDWAG 360
Db	318	DFDALKDIQNLKEKIFAIEGTETITSSSSFELEWAQEGFSAVFTPDGVLGAVGFWWSG 377
QY	361	GVFLYTSKEKSTFNMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420
Db	378	GAFLYPPNMSPTFINMSQENVMDRDSVILGYSTELALWKGVQSLVLGAPRYOHTKAVIFT 437
QY	421	QNTGQWESNANVKCTOICGAVGASLCSVDVDSNGSTDLVLIGAPHYVQTRGGQVVCPL 480
Db	438	QVSWQWMAEVTQTGIVGFGSLCSVDVDTGSDTLVLIGAPHYVQTRGGQVVCPL 497
QY	481	PRGORARQCDVAVLYGQGPQWGRFGAALTVLGVDWNGDKLTDVAIGAPGEDNAGAVLYF 540
Db	498	PRGWR-RWMCDAVLYGQGHQWGRFGAALTVLGVDWNGDKLTDVVGAPGEENRAGAVLYF 556
QY	541	HGTSGSISPSHSQRIAGSKLSPLOVFGQSLSGQDLTMDGLVDLTGAGQGHVLLRSQ 600
Db	557	HGVLGPISSPSHSQRIAGSQLSSRLQVFGQALSGQDLTQDGLVDLAVGARGQVLLLR 616
QY	601	PVLVRVKAIMEPNREAVRNVEPCNDVYVKGAEVRCVCHLVQKSTDRDLREGIOISVVT 660
Db	617	PVLVWGVSMQFIPAEIRPSAFECREQVVBQTLVQSNICLYIDRKNLLGSRDLQSVT 676
QY	661	YDLALDSGRPHSRVAFNETKNSTRQQTQVLGLTQTCETKLQLEPNCIEDPVSPVTLRLNF 720
Db	677	LDLALDPGRLSPRATFOETKNRSLSRVRLGLKAHCENFLLLPSCVEDSVTPITLRLNF 736
QY	721	SLVGTPLSAFGNLRPVLAEDAQRILFTALFPFEKNCNDNICODDLSTFTFSMLDCLVVG 780
Db	737	TLVGKPLLAFLNLRPMLAALQAQRYFTASLPFEKNCAGADHICODNLGIFSFPGLKSLVG 796
QY	781	GPREFNVTVVRNDGEDSVYQVTFEFPDLVSRKVTQLQNRQSQRWRACBSASSTEV 840
Db	797	SNLELNEVWMDGEDSVYGTITFSPAGLSYRYVAEGQKQGLRSLHLTCDSPVYG-- 854
QY	841	SGALKSTSCSINHIPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTFF 900
Db	855	SQGTWSTCRINHLIFRGAQITFLATFDVSPKAVLGDRLLLLTANVSSENNTPRTSKTTP 914
QY	901	QLELPVKYAVMYVTSHGVTXYLNFTAS-ENTSRVMOHQVQVSNLQORSLPISLVLVP 959
Db	915	QLELPVKYAVYTVVSSHEQTKYLNFSSEKESHVAMHYQVNNLQORDLPVSINFWVP 974
QY	960	VLNQTWIDRPOVTFSENSSITCHTERLPSSHDFLAELRKAPVNCSTAVQRCQCDI 1019
Db	975	VELNQEAVMWDMVESHQNPQNSCSSEKIAAPPASDFLAHQKNEPVLDCSIAGCLRFRCDV 1034
QY	1020	PFFGIQIEFNATLKGNSLFDWYIKTSHNHLILVSTAILFNDSVFTLLPGQGAFFVRSQTE 1079
Db	1035	PSFSVQBELDTLKGNSLFGWVRQILQKKVSVVSAEITFTDTSVYSQLPGQEAFFRAQTT 1094
QY	1080	TKVPEFVEPNPLPIVSSVGLLLALITAAALYKLGFFKQYKDMSE 1128
Db	1095	TVLEKYKVHNPTPLVSSIGLLLLALITAVLYKVGFQYKEMMEE 1143
RESULT 13		
AAAR07120		
ID	AAAR07120	standard; protein; 1163 AA.
XX	AC	AAAR07120;
XX	DT	25-MAR-2003 (revised)
DT	05-FEB-1991	(first entry)
XX	DE	p150.95 alpha subunit encoded by clone lambda47.
XX	KW	p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia;
KW	rhinovirus.	
XX	OS	Synthetic.
XX	Key	Location/Qualifiers
FT	Region	1..19
FT	Region	/label= signal peptide
FT	Region	20..44
FT	Modified-site	/label= N-terminus
FT	Modified-site	61..63
FT	Modified-site	/label= glycosylation site
FT	Modified-site	89..91
FT	Modified-site	/label= glycosylation site
FT	Modified-site	385..387
FT	Modified-site	/label= glycosylation site
FT	Modified-site	392..394
FT	Modified-site	/label= glycosylation site
FT	Modified-site	697..699
FT	Modified-site	/label= glycosylation site
FT	Modified-site	735..737
FT	Modified-site	/label= glycosylation site
FT	Modified-site	899..901
FT	Modified-site	/label= glycosylation site
FT	Modified-site	904..906
FT	Modified-site	/label= glycosylation site
FT	Modified-site	939..941
FT	Modified-site	/label= glycosylation site
FT	Modified-site	1050..1052
FT	Modified-site	/label= glycosylation site
FT	Domain	1108..1133
FT	Domain	/label= transmembrane
XX	WO9010646-A.	
PN	20-SEP-1990.	
XX	09-MAR-1990;	90WO-US001257.
XX	09-MAR-1990;	90WO-US001257.
XX	(DAND )	DANA FARBER CANCER INST INC.
PI	Corbi AA,	Springer TA;
XX	WPI;	1990-304985/40.
DR	N-PSDB;	AAQ06068.
XX	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-	
PT	unit of p150.95 cell surface adhesion receptor, opt. together with a beta	
PT	chain of CD-18 family.	
XX	Disclosure; Fig 3; 59pp; English.	

CC Clone lambda X47 was isolated from a cDNA library constructed from total  
CC RNA extracted from phorbol myristate acetate stimulated HL-60  
CC myelomonocytic cells. The library was screened with oligonucleotide  
CC probes based on tryptic peptide fragments of p150.95. The sequence can be  
CC attached to appropriate control elements and expressed in prokaryotic and  
CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral  
CC infection because it interacts with ICAM-1 and inhibits cell-virus  
CC attachment. It can also be used as an anti-inflammatory agent. See also  
CC AAQ06083-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003  
CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)  
XX  
SQ Sequence 1163 AA;

Query Match 58.5%; Score 3437; DB 2; Length 1163;  
Best Local Similarity 60.3%; Pred. No. 3e-275;  
Matches 681; Conservative 143; Mismatches 299; Indels 6; Gaps 4;  
SQ  
QY 1 FNLDTENAMTFOENARGFGQSYYVQLQGRVVVVGAPQEIYAANQSGSLYQCDYSTGSCBPI 60  
DB 20 FNLDTEELTAFRVDSAGFGDSVVQYANSVVVVGAPQKITAANQTGGLYQCGYSTGACBPI 79  
QY 61 RLQVPEAVNMSLGLSLAATSPDQLLACGPTVHOTCSENTYVKGCLFLFSGNLRLQOQOK 120  
DB 80 GLQVPEAVNMSLGLSLASTTSPSOLLACGPTVHHECGENMYLTLGLCFLGPT--QLTQR 137  
QY 121 PPEARLGCPOEDSDIAFLIDGSGSIIPHDFRKMELVSTIMEQLKSKTLFSLMOYSBEF 180  
DB 138 LPVSQOECPQODIVFLIDGSGSISSRNFAFMNPFRAVTSQFORPTQFSLMGFSNKF 197  
QY 181 RIHFTFKFQNNPNRSLKPTQLILGRTHRTATGLRKVVRLFNITNGARKNAKILFLL 240  
DB 198 QTHFTFEPRRTSNPLSLASVHQLQGGFTYTATATQNVHRLFHASYGARRDATKILIVI 257  
QY 241 TDGKFGDPLGYEDVIPLEDRGVIRYVLGFGDAPRSEKSRQELNVTASKPRDHVQAN 300  
DB 258 TDGKREGSLDYKDVIPMAADAALIRYAIQVGLAFQNRNSWKELDIAKPSQEHIFKVE 317  
QY 301 NFEALKTWONLREKIFATEGTQGTGSSSFEHEMGEFSAATTSNGPLLSVTGSDWAG 360  
DB 318 DFDALKDIONLKEKIFALEGTETSSSFELEMAQGFSAVFTPDGVLGAVGFTWSG 377  
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILRNVRQSLVLGAPRYQHIGLVAMPR 420  
DB 378 GAFLYPPNMSPIFINMSQENVDMDRDSYLGSTELALWKGVSILVGLAPRYQHTGKAVFT 437  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 480  
DB 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYEQTRGGQVSVCP 497  
QY 481 PRGQARMOCDAVLVGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540  
DB 498 PRGWR-RWCDVLYGEQHPWGRFGAALTVLGDVNGDKLTDVVAIGAPGEEDNRGAVLYF 556  
QY 541 HGTSGSGISPSHSORISAGSLSPRLQYFGQSLSGGQDLMGDLVLTGAGQHVLLRSQ 600  
DB 557 HGVLFPSISPSHSORISAGSLSPRLQYFGQSLSGGQDLMGDLVLTGAGQHVLLRSR 616  
QY 601 PVLRVKAIMFENPREVARNVFNCDQVVGKEAGEVRVCLHVQKSTRDLREGQIOSVT 660  
DB 617 PVLWVGVSQFTPAEIPRSAFECEQVSEQTILVQSNICLYIDKRKSNLLSGRDLQSSVT 676  
QY 661 YDLALDSGRPHSAVENETKNSFRTOVLGLTQTCETLKLQLPNCIEDPVPISVLRLNF 720  
DB 677 LDALDPLRLSPRATQETKNSRLSRVULGLKAHCENFLLLPVCSVSVPITLRLNF 736  
QY 721 SLVGTPLSAFNLRLPVLAEADQRLFTALFFPKNCGNDNICDDLSITFSMSLDCLVVG 780  
DB 737 TLVGLKLLAFNLRLPMLAALQRYFTASLFPFNKCGADHICQDNLGIFSFPGLKSLVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPLDLSYRKYSTONQORSQSWRLACESASTEV 840  
DB 797 SNLELNAEVNVMVNDGEDSYRTTTFPSHAPGLSYRYVABGQKQGLRSLHLTCDSA--PVA 854

QY 841 SGALKSTSCSINHPIFFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTKTEF 900  
DB 855 SQGTWSTSCRINHILFRGGAQITFLATFDVSPKAVLGRLLLTANVSSENTPRTSKTTF 914  
QY 901 QLELPVKYAVNVTSHGVSTKYLNFATAS-ENTSRVMOHQVSVNLGORSPLISLVELVP 959  
DB 915 QLELPVKYAVTVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLGORDLPVSNFWFP 974  
QY 960 VRLNQVTWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDI 1019  
DB 975 VELNQBAVMMDEVSHPQNPRLRSCSSQKIAPPASDFLAHIQKNPVLDCSIAGCLFRCDV 1034  
QY 1020 PFFGIQEEFNATLKNLSFDMWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTE 1079  
DB 1035 PFSVQEBELDTLKNLSFGWVRQILOKKVSVSVVAEITDTSVTSQLPQGERFMAQTT 1094  
QY 1080 TKVPEFVNPDLPLIVGSSVGLLLALLITAAALYKLGFFKQYKDMNSE 1128  
DB 1095 TVLEKYKVHNPPLIVGSSIGALLLALLITAVLYKVGFQYKEMMBE 1143  
RESULT 14  
AAW65091  
ID AAW65091 standard; protein; 1163 AA.  
XX  
AC AAW65091;  
XX  
DT 28-SEP-1998 (first entry)  
XX  
DE Human Beta-integrin CD11c subunit protein.  
XX  
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;  
KW rheumatoid arthritis.  
XX  
OS Homo sapiens.  
XX  
PN US5728533-A.  
XX  
PD 17-MAR-1998.  
XX  
PF 07-JUN-1995; 95US-00485618.  
XX  
PR 23-DEC-1993; 93US-00173497.  
PR 05-AUG-1994; 94US-00286889.  
PR 21-DEC-1994; 94US-00326652.  
XX  
PA (ICOS-) ICOS CORP.  
XX  
PI Van Der Vieren M, Gallatin WM;  
XX  
XX WPI; 1998-206565/18.  
DR  
XX Screening assay for modulators of integrin binding - using immobilised or  
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.  
XX  
PS Example 5; Fig 1A-D; 106pp; English.  
XX  
XX This sequence represents a human beta-integrin CD11c subunit which is  
CC used to describe a method for identifying compounds that modulate the  
CC interaction of the beta-integrin alpha-d subunit with a binding partner  
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
CC -d binding partner, one of which is immobilised and the other of which is  
CC labelled, in the presence of a test compound, and determining if the  
CC compound affects binding between the alpha-d polypeptide and alpha-d  
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
CC comprising the cytoplasmic, transmembrane or extracellular domain of  
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat  
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
CC and rheumatoid arthritis  
XX

SQ Sequence 1163 AA;  
 Query Match 58.2%; Score 3423; DB 2; Length 1163;  
 Best Local Similarity 60.2%; Pred. No. 4.4e-274;  
 Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4

QY	1	PNLTENAMTQENARGFGQSVVLGGRRVVGAPQETVAANQRGSLYOCYSTGSCBPI	60
DB	20	FNLDEELTAFRVDSAGFGSDVQVANSVWVVGAPQKIAANQIGGLYOCGYSTGACBPI	79
QY	61	RLQVPEAVANMSLGLSLAATTPPOLACGPTVHCTCSENTYVKGICLFLGSLNBOQPK	120
DB	80	GLQVPEAVANMSLGLSLAATTPPOLACGPTVHCTCSENTYVKGICLFLGSLNBOQPK	137
QY	121	FPEALRCPOEDSDIAFLIDGSGSTIIPHDFRMKELVSTIMEQLKSKTFLSLMOYSEF	180
DB	138	LPVSRQCPREQDQIVFLIDGSGSISSRFATMNFVRAVISQFQPTQSLMWFQSKNF	197
QY	181	RIIHFPKFEQNNPNSRLIKPITQLGRTHATGLRKVVRELFNITNGARKNAFKILFL	240
DB	198	QTHFTFEFRRTSNPLSILASVHQLQGTYTATAQNVVHRLFHASYGARRDAIKILVI	257
QY	241	TDBEKFGDPLGYEDVIPEDOREGIVRYVLGFGDAFRSEKSRQELINTVASKPRDHVFOAN	300
DB	258	TDGKKGSDLDYKDVIPMAADAGIIRVAIGVLAFQNNRNSWKELEIDASKSQSHIFKE	317
QY	301	NFEALKTVQNLREKIFAIEGTQGTGSSSFHEMSQEGFSAATISNGPLLTSTVGSYDWAG	360
DB	318	DFDALKDIQNLKEKIFAIEGTETIISSSFLEWAQEGFSVTFDGPVLGAVGSGFTWSG	377
QY	361	GVFLYTSKSKSTFNNTRVDSDDNDAVILGYAAAIIILNRVOSLVLGAPRYQHIGLVAMFR	420
DB	378	GAFLYPPNNMPTTFINMSQENVDMRDSYLGSTELALWKGVSQSLVGLAPRYQHIGKAVIFI	437
QY	421	QNTGMWESNANVKTQIGAVFGASLCSDVDNSGSTDVLVILGAPHYYEQTGGQSVSCPL	480
DB	438	QVSRQWPKAEVIGTQIGSTFGSLCSVDVDTGSTDVLVILGAPHYYEQTGGQSVSCPL	497
QY	481	PRGQARWQCDADVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF	540
DB	498	PRGWR-RWMCADVLYGEGQHPWGRFGAALTVLGVDVNGDKLTDVVI GAPGEENRGAVYLF	556
QY	541	HGTSGSGISPSHSGRIAGSKLSPRYFGQSLGGQDLTMDGLVDLTVGAGQHVLLLRSQ	600
DB	557	HGVLGPIPSPSHSGRIAGSKLSPRYFGQSLGGQDLTMDGLVDLVAVGARGQVLLLR	616
QY	601	PVLVRKAIMEFNPREVARNVFECDNDVVVKEAGEVRVCLHVQKSTRDLRBEQIQSVVT	660
DB	617	PVLVVGVSQMFIAEIPRFAFEQREQVVSQTLVQSNICILYDKRKNLLGSLRDLQSSVT	676
QY	661	YDALDSGRPHSRVAFNETKNSTRRTQVLLGTQTCTLKLQLPNCIEDPVSPTVLRLNF	720
DB	677	LDLALAPGRLSPRAIQETKNRSLSRVVLGLKAHCENFNLPLSPCEDSVIPIILRLNF	736
QY	721	SLVGTPLSAPGNLRPVLAEDAQRFLTALFPFEKNCGNDNICQDDLSITFSFMSLDCLVWG	780
DB	737	TLVGKPLIAPRNLPRMLAALQAQRYFTASLPFEKNCGADHICQDNLGISFSFPFGKLKLVG	796
QY	781	GREPEFVTVTRNDGEDSYRTQTFPPPLDLVSRKVVSTLQNRQSRWRVLRACESASTEV	840
DB	797	SNUELNAEVMVWMDGEDSYGTTITFSPAGLSRYVAEGQKQGLRSLHITC--CSAPVG	854
QY	841	SGALKSTSCSINHPIFPENGEVTFNITFDVDSKASLGKLLKANVTSENMMPTNKTFF	900
DB	855	SOQTWSTSCINHLIFPGGAQIITFLATFDVSPKAVGLDRLILLIANVSENNIIPRTSKTIF	914
QY	901	QLELPVKYAYVMVVTSHGVSTKYLNFITAS-ENTSRVMQHOYQVSNLQGRSLIPSLVFLVP	959
DB	915	QLELPVKYAYIVVSSHEQTKYLNFSSEKESHVAMHRYQVNNLQORDLPVSIINFWVP	974
QY	960	YRLNQIVMDRPOVITSENLSSTCHTKERIIPSHDSFLAELRKAPVNVNCSIAVQRIQCDI	1019
DB	975	VELNQEAVMMDVRSHPQNPRLSCSEKSIAPPASDFLAHIQKNPVLDCSLTAGCIRFCDV	1034

QY	1020	PFFGIGEEFNATLKGNI	SFDWYIKTSNNHLLIVSTAEILNDSVFTLLPGGAFVRSQTE	1079
		:	:	
Db	1035	PSFSVQBELDFTLKGNI	SFGWVRLOIKQKSVSVSVAEIIIDTSVYSQLPGEAFMRAQTI	1094
		:	:	
QY	1080	TKVBSFEVPNPPLPIV	GSSVGGELLALITAAALYKLGFFKROYKDMWSE	1128
		:	:	
Db	1095	TVLEKYKVHNPPIV	GSSVGGELLALITAVLYKVGFFKROYKEMWEE	1143
		:	:	
RESULT 15				
AAB07361				
ID	AAB07361	standard; protein; 1163 AA.		
XX				
AC	AAB07361;			
XX				
DT	21-JAN-2001	(first entry)		
XX				
DE	Human CD11c protein sequence.			
XX				
KW	Human; macrophage infiltration inhibition; alpha_d integrin;			
KW	leukocyte integrin; Leu-CAM; leukointegrin; immune response;			
KW	inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;			
KW	atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;			
KW	lung inflammation; acute respiratory distress syndrome; Crohn's disease;			
KW	rheumatoid arthritis; central nervous system injury; CD11c.			
OS	Homo sapiens.			
XX				
PN	WO200029446-A1.			
XX				
PD	25-MAY-2000.			
XX				
EF	16-NOV-1999; 99WO-US027139.			
XX				
PR	16-NOV-1998; 98US-00193043.			
PR	08-JUL-1999; 99US-00350259.			
XX				
PA	(ICOS-) ICOS CORP.			
XX				
PI	Gallatin MW, Van Der Vieren M;			
XX				
DR	WPI; 2000-387751/33.			
XX				
PT	Use of novel anti-alpha integrin d monoclonal antibodies to inhibit			
PT	macrophage infiltration and reduce inflammation at central nervous system			
PT	injury sites.			
XX				
PS	Example 5; Fig 1; 270pp; English.			
XX				
CC	Integrins are a class of membrane-associated molecules that participate			
CC	in cellular adhesion. Integrins are made up of an alpha subunit and a			
CC	beta subunit. One class of human integrins are restricted to expression			
CC	in white blood cells and have a common beta2 subunit: the leukocyte			
CC	integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins			
CC	have an important role in immune and inflammatory responses. The present			
CC	protein sequence is the human integrin alpha subunit CD11c. This sequence			
CC	was used in an alignment to identify a novel beta2 integrin alpha			
CC	subunit: alpha_d (AA060014 and AAB07359). The present sequence has			
CC	approximately 66% identity to the protein sequence of alpha_d. The			
CC	Alpha d gene and protein may be useful in therapy for diseases linked to			
CC	aberrant alpha_d function e.g. Type I diabetes, atherosclerosis, multiple			
CC	sclerosis, asthma, psoriasis, lung inflammation, acute respiratory			
CC	distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency			
CC	(LAD). In addition, anti-alpha_d monoclonal antibodies may be used in the			
CC	inhibition of macrophage infiltration at the site of a central nervous			
CC	system injury. The monoclonal antibodies can also be used to detect and			
CC	diagnose Crohn's disease			
XX				
SQ	Sequence 1163 AA;			

Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;			
Qy	1	FNLDTENAMTFOENARGFGQSVVQLOGSRVVVVGAPQEIIVAAQNGSLIQCYSYSGSCEPI	60
Db	20	FNLDTTEELTAFRVDSAGFGSDVVQVANSVVVVGAPQKIIAANQIGGLYQCGYSTGACEPI	79
Qy	61	PLQVPEAVNMSLGLSLAATSPPOLACGPTVHQTCSNTYVKGCLFLFGSLNLRQOPQK	120
Db	80	GLQVPEAVNMSLGLSLAATSPPOLACGPTVHHCGRNMYLTGLCLFLGPT--QLTQR	137
Qy	121	PPEALRGCPQSDSDIAFLIDSGSIIIPHDFFRMKELVSTIMEQLKKSKTLFSLMQYSEEP	180
Db	138	LPVSRQECPRQEQDIVFLIDSGSISRRNFATMNFVRAVISQFORPSTQFSLMQFSNKF	197
Qy	181	RIHFTFKFQNNPNRPSIKPITQLLGRTHATGLRKVVURELFNITNGARKNAFKILFLL	240
Db	198	QTHFTFEFRRTSNPLSLASVHQLQGFTYTATATQNVVHRLFHASYGARRDAIKILIVI	257
Qy	241	TDGEKFGDPLGYEDVPELDRGVIRYVLFGDFAFRSEKSRQELNVTASKPPDRDHVFOAN	300
Db	258	TDGKKEGSDLYKOVIPWADAAGIIRYAIQVGLAFQNRNSWKELNDIASKPSQEHIFKVE	317
Qy	301	NPEALKTVQNOLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITNSGPLLSTVSGYDWAG	360
Db	318	DFDALKDQONQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTPDGPVLGAVGSFTWSG	377
Qy	361	GVFLYTSKEKSTFINMTRVDSQMDNDAYLGYAAAIILNRNVQSLVLGAPRYOHIGIVAMPR	420
Db	378	GAFLYPPNMSPTFINNSQENVMDRDSYLGSTELALWKGVQSLVLGAPRYOHIGKAVIFI	437
Qy	421	QNTGMWESNANVKGFTQIGAYFCASLCSVDVDSNGSTDVLIGAPHYETRGQVSVCP	480
Db	438	QVSRQWRMAEVIQIGSYFQASLCSVDVDTGSDTDLVLIGAPHYETRGQVSVCP	497
Qy	481	PRGQARWOCDAVLXGEQCPGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVILF	540
Db	498	PRGWR--RWCDAVLYGEQGHMGRFGAALTVLGDVNGDKLTDVWIGAPGEENRGAVILF	556
Qy	541	HGTSSGSIQSPSHQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVCAGHVLILRSQ	600
Db	557	HGVLGFSISPSHSQRIAGSKLSPRLQYFGQALSGGQDLTQDGLVDLVAVARGQVLLRTR	616
Qy	601	PVLRVKAIIEENPREVARVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
Db	617	PVLWVGVSQCFPAEIPRAPECREQVSEQTLVQSNICLYDKRSKNLLGSRDLQSSVT	676
Qy	661	YDLALDSGRPHSAFNETKNSTRQTQVLGTQTCETLKLQLPNCIEDPVPVILRLNF	720
Db	677	LDLALAPGRLSPRAIFOETKNRSLSRVRLGLKAHCENFNLLLPSCVEDSVIPIILRLNF	736
Qy	721	SLVGTPLSAFGLNRPVLAEDAQPLFTALPFEKNCNDNITCODDLSITFSEMSLDCLVYG	780
Db	737	TLVGKPELLAFRLNRPMLAALQRYFTASLPFEKNCADHICQDNLGISFSFGLKSLLVG	796
Qy	781	GPREFNVTVRNDRGDSYRTOVTFPPFLDLSVRKYSTLQNRQSRQSWRLACESASSTEV	840
Db	797	SNLELNAEVMVNDGEDSVGTITTFSHPAGLSRYRYVABGQQLRSLHLTC--CSAPVG	854
Qy	841	SGALKSTCSINHIPIFENSEVTFTNITFDVDSKASLGNKILLKANVTSENMPRTNKTFF	900
Db	855	SQGTWSTSCRINHLIFRGGAQITFLATFDVSPKAVGLDRLLLIANTVSSENNIPTSKTIF	914
Qy	901	QLELPVKYAVYVTVSHGVSTKVLNFTAS--ENTSRVWQHGYQVSNLQORSLPISLVFLYP	959
Db	915	QLELPVKYAVYIVVSSHEQFTKYNFSESEKESHVAMHRYQVNNLQGRDLPVSINFVWP	974
Qy	960	VRNLQTVIWRDPQVTFSENLSSTCHTKERLPDLSFLAELRKAPVWVNCIAVCQRIQCQDI	1019
Db	975	VELNQAVWMDVESHVPQNPQLRCSSEKIPAPASDFLAHQKNPVLDCSIAGCLRFECDV	1034
Qy	1020	PFFGIOEEFNATLKNLSFDWYIKTSHNLLIVSTAEILPNDVSTFLLPQOGAFVRSQTE	1079
Db	1035	PSFSVQBEELDTLKGNLISFGWVRQILOKQKSVSVSVAEIIIPDTSVYSQLPQGEAFMRAQTI	1094

Qy 1080 TKVEPFEVENPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSE 1128  
Db 1095 TVLEKYKHNPILPLIVGSSIGGLLLALITAVLYKVGFKKQYKEMMEE 1143

Search completed: November 9, 2004, 12:08:14  
Job time : 109.25 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds  
(without alignments)  
4119.157 Million cell updates/sec

Title: US-09-902-481B-3  
Perfect score: 5879  
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMWSEGGPGEAPQ 1137

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09D\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5879	100.0	1137	10	US-09-902-481A-3
2	5857	99.6	1137	10	US-09-902-481A-4
3	5839	99.3	1153	9	US-09-350-259-3
4	5839	99.3	1153	10	US-09-902-481A-1
5	5839	99.3	1153	10	US-09-891-943-3
6	5839	99.3	1153	14	US-10-144-259-30
7	5839	99.3	1153	14	US-10-207-655-176
8	5836	99.3	1137	10	US-09-902-481A-5
9	5832	99.1	1137	10	US-09-902-481A-6
10	5823.5	99.1	1152	9	US-09-945-265-4
11	3446	58.6	1163	14	US-10-116-275-204
12	3423	58.2	1163	9	US-09-350-259-4
13	3423	58.2	1163	10	US-09-891-943-4

14	3388	57.6	1161	9	US-09-350-259-2
15	3388	57.6	1161	10	US-09-891-943-2
16	3372.5	57.4	1161	9	US-09-350-259-99
17	3372.5	57.4	1161	10	US-09-891-943-99
18	3213.5	54.7	1161	9	US-09-350-259-55
19	3213.5	54.7	1161	10	US-09-891-943-55
20	3201	54.4	1161	9	US-09-350-259-53
21	3201	54.4	1161	10	US-09-891-943-53
22	3192.5	54.3	1151	9	US-09-350-259-37
23	3192.5	54.3	1151	10	US-09-891-943-37
24	3180	54.1	1155	9	US-09-350-259-46
25	3180	54.1	1155	10	US-09-891-943-46
26	1848	31.4	369	13	US-10-087-192-1212
27	1536.5	26.1	1170	9	US-09-945-265-2
28	1532.5	26.1	1170	15	US-10-261-164-1
29	1509	25.7	1223	16	US-10-408-765A-295
30	1341.5	22.8	1086	16	US-10-408-765A-1871
31	1229.5	20.9	494	9	US-09-350-259-103
32	1229.5	20.9	494	10	US-09-891-943-103
33	1128.5	19.2	413	9	US-09-350-259-101
34	1128.5	19.2	413	10	US-09-891-943-101
35	1128	19.2	1179	14	US-10-173-550-2
36	1128	19.2	1179	14	US-10-173-551-2
37	1087	18.5	1188	14	US-10-291-265-810
38	1085.5	18.5	1151	10	US-09-884-130-103
39	1085.5	18.5	1151	10	US-09-836-353A-103
40	1085.5	18.5	1179	10	US-09-918-715-250
41	1085.5	18.5	1179	17	US-10-474-794-250
42	1083	18.4	1188	14	US-10-291-265-338
43	1079.5	18.4	1189	10	US-09-884-130-35
44	1079.5	18.4	1189	10	US-09-836-353A-35
45	1079.5	18.4	1189	15	US-10-262-839-4

## ALIGNMENTS

## RESULT 1

US-09-902-481A-3  
; Sequence 3, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-3

Query Match 100.0%; Score 5879; DB 10; Length 1137;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1137; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANQSGSLYQCDYSTGSCPT	60
Db	1	FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANQSGSLYQCDYSTGSCPT	60
Qy	61	RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHOTCSNTYVKGCLFLGSLNLRQPOK	120
Db	61	RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHOTCSNTYVKGCLFLGSLNLRQPOK	120



; Publication No. US20030054440A1									
; GENERAL INFORMATION:									
; APPLICANT: Springer, Timothy									
; APPLICANT: Shimooka, Motomu									
; APPLICANT: Shifman, Julia									
; APPLICANT: Mayo, Stephen									
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY									
; FILE REFERENCE: A-70586-1/RT/RMS/RMK									
; CURRENT APPLICATION NUMBER: US/09/902,481A									
; CURRENT FILING DATE: 2001-07-09									
; PRIOR APPLICATION NUMBER: US 60/216,600									
; PRIOR FILING DATE: 2000-07-07									
; NUMBER OF SEQ ID NOS: 7									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 4									
; LENGTH: 1137									
; TYPE: PRT									
; ORGANISM: Artificial sequence									
; FEATURE:									
; OTHER INFORMATION: synthetic									
US-09-902-481A-4									
Query Match 99.6%; Score 5857; DB 10; Length 1137;									
Best Local Similarity 99.5%; Pred. No. 0;									
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;									
QY	1	FNLDTENAMTFOENARGFGQSVVQLQGGRRVVVGAPQEI	VAANQSGSLYQCDYSTGSCPEI	60					
DB	1	FNLDTENAMTFOENARGFGQSVVQLQGGRRVVVGAPQEI	VAANQSGSLYQCDYSTGSCPEI	60					
QY	61	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTYVKGICFLFGNLRQPOK	120						
DB	61	RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTCSENTYVKGICFLFGNLRQPOK	120						
QY	121	FPEALRGCPQEDSDTAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLESI	MOYSEEF	180					
DB	121	FPEALRGCPQEDSDTAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTLESI	MOYSEEF	180					
QY	181	RHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFKILFLL	240						
DB	181	RHFTFKFQNNPNRSLIKPITQLLGRTHRTATGLRKVVRELFNITNGARKNAFKILFLL	240						
QY	241	TDGEKFGDPLGYEDVIPELDREGVIRYVLGFDGDAFRSEKSRQELNTVASKPPRDHVFQAN	300						
DB	241	TDGEKFGDPLGYEDVIPELDREGVIRYVLGFDGDAFRSEKSRQELNTVASKPPRDHVFQIN	300						
QY	301	NFEALKTIVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITTSNGPLLSTVGSYDWAG	360						
DB	301	NFEALKTIVQNLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITTSNGPLLSTVGSYDWAG	360						
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAIILRNVRQSVILGAPRYQHIGLVAMER	420						
DB	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAIAIILRNVRQSVILGAPRYQHIGLVAMER	420						
QY	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYEQTRGGQSVCPPL	480						
DB	421	QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVIGAPHYEQTRGGQSVCPPL	480						
QY	481	PRQQRARWQCDVLYGEGQPGWGRFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAVYLF	540						
DB	481	PRQQRARWQCDVLYGEGQPGWGRFGAALTIVLDVNGDKLTDVAIGAPGEDNRGAVYLF	540						
QY	541	HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ	600						
DB	541	HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ	600						
QY	601	PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTDRREGQIQSVVT	660						
DB	601	PVLVRKAIIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTDRREGQIQSVVT	660						
QY	661	YDLALDSGRPHSRAVFNETKNSTRQTVGLGTQTCETLKLQLPNCIEDPVPSPVLRINF	720						
DB	661	YDLALDSGRPHSRAVFNETKNSTRQTVGLGTQTCETLKLQLPNCIEDPVPSPVLRINF	720						

```
Qy 721 SLVGTPLSAFGLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 721 SLVGTPLSAFGLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Qy 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Qy 901 QLELPVKYAVMVVTSYKYNFTASENTSRVMOHQYVSNLQORSLSLVLVLPV 960
Db 901 QLELPVKYAVMVVTSYKYNFTASENTSRVMOHQYVSNLQORSLSLVLVLPV 960
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
Db 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
Qy 1021 FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
Db 1021 FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
Qy 1081 KVEPFEVNPPLPIVSGSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
Db 1081 KVEPFEVNPPLPIVSGSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137

RESULT 3
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 9; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGSVVLQGSRRVVVGAPEIVAANQRGSIYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTGLSLAAATTPPOLLACGPTVHQTCSNTYKGLCFGLGSNLRQOPQK 136
Qy 61 RLQVPVEAVNMSGLSLAAATTPPOLLACGPTVHQTCSNTYKGLCFGLGSNLRQOPQK 120
Db 77 RLQVPVEAVNMSGLSLAAATTPPOLLACGPTVHQTCSNTYKGLCFGLGSNLRQOPQK 136
Qy 121 FPEALRCGPQEDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTILFSLMOYSEEP 180
Db 121 FPEALRCGPQEDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTILFSLMOYSEEP 180
```

```
Db 137 FPEALRCGPQEDSDIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTILFSLMOYSEEP 196
Qy 181 RIHFTTFKEFQNNPNRSLIKPIITQLLGRTHTATGLRKVVRLEFNITNGARKNAFKILFLL 240
Db 197 RIHFTTFKEFQNNPNRSLIKPIITQLLGRTHTATGLRKVVRLEFNITNGARKNAFKILVVI 256
Qy 241 TDGKFGDPLGYEDVPELDRGVIRVVLGFGDAFRSEKSOELNTVASKPRDHVFOAN 300
Db 257 TDGKFGDPLGYEDVPELDRGVIRVVLGFGDAFRSEKSOELNTVASKPRDHVFOAN 316
Qy 301 NFEALKTVQNLREKIFAIEGTQTGSSSSSEHEHMSQGFSAAITSNGLLSTVGSYDWAG 360
Db 317 NFEALKTVQNLREKIFAIEGTQTGSSSSSEHEHMSQGFSAAITSNGLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496
Qy 481 PRGQARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRGQARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 556
Qy 541 HGTSGSISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLLLRSQ 600
Db 557 HGTSGSISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGHVLLLRSQ 616
Qy 601 PVLRVKAIMBPNPREVARNVFECDNVKKEAGEVRVCLHVQKSTDRLEGIOQSVVT 660
Db 617 PVLRVKAIMBPNPREVARNVFECDNVKKEAGEVRVCLHVQKSTDRLEGIOQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQNPCTEDPVSPTVRLNFP 720
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQNPCTEDPVSPTVRLNFP 736
Qy 721 SLVGTPLSAFGLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 840
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSWRLACESASSTEV 856
Qy 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
Qy 901 QLELPVKYAVMVVTSYKYNFTASENTSRVMOHQYVSNLQORSLSLVLVLPV 960
Db 917 QLELPVKYAVMVVTSYKYNFTASENTSRVMOHQYVSNLQORSLSLVLVLPV 976
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1036
Qy 1021 FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080
Db 1037 FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1096
Qy 1081 KVEPFEVNPPLPIVSGSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1137
Db 1097 KVEPFEVNPPLPIVSGSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPQ 1153

RESULT 4
US-09-902-481A-1
; Sequence 1, Application US/09902481A
; Publication No. US2003005440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
```

```
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match          99.3%; Score 5839; DB 10; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSLNRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSLNRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKTLFSLMQYSEEF 196

QY 181 RHFTFKFQNNPNRSLIKPTTQLGRTHATGLRKVVRELFTNIINGARKNAFKILFL 240
Db 197 RHFTFKFQNNPNRSLIKPTTQLGRTHATGLRKVVRELFTNIINGARKNAFKILFL 256

QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
Db 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 316

QY 301 NPEALKTVONQUREKIFAIEGTQTGSSSFHEHMSQEGFSAITSNGLPLSTVGSYDNAG 360
Db 317 NPEALKTVONQUREKIFAIEGTQTGSSSFHEHMSQEGFSAITSNGLPLSTVGSYDNAG 376

QY 361 GVFLYTSKESKTFINMTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKTFINMTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVVCP 496

QY 481 PRGQARWQCDVLYEQQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVILF 540
Db 497 PRGQARWQCDVLYEQQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVILF 556

QY 541 HGTSGSGISPSHSORITAGSKLSPLQYFGQSLSGGQDLTMDGLVLDITVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORITAGSKLSPLQYFGQSLSGGQDLTMDGLVLDITVGAQGHVLLRSQ 616

QY 601 PVLVRKAIAMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 660
Db 617 PVLVRKAIAMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 676

QY 661 YDLALDSGRPHSAVENETKNSRTRQTVLGLTQCTETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSAVENETKNSRTRQTVLGLTQCTETLKLQLPNCIEDPVSPIVLRNLF 736

QY 721 SLVGTPLSAFGNLRPVLAEDAQRULTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGNLRPVLAEDAQRULTALFPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796

QY 781 GPRFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRLACESASSTEV 840
Db 797 GPRFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFFPENSEVTENITFDVDSKASGLNKLKLLKANVTSENMPRNKTEF 900
Db 857 SGALKSTSCSINHPIFFPENSEVTENITFDVDSKASGLNKLKLLKANVTSENMPRNKTEF 916

QY 901 QLELPVKYAVVMVVTSHGVSTKYLNTASENTSRVMQHQYQVSNLQGSRLPLSLVFLVPV 960
Db 917 QLELPVKYAVVMVVTSHGVSTKYLNTASENTSRVMQHQYQVSNLQGSRLPLSLVFLVPV 976

QY 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPHSDFLAELRKAPVTVNCISIAVCORTODIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSTCHTKERLPHSDFLAELRKAPVTVNCISIAVCORTODIP 1036

QY 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDVSFTLLPQOGAFVRSQTET 1080
Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEIILFNDVSFTLLPQOGAFVRSQTET 1096

QY 1081 KVEPFEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMMSGGPPGABPQ 1137
Db 1097 KVEPFEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMMSGGPPGABPQ 1153

RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match          99.3%; Score 5839; DB 10; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSLNRQOPQK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSLNRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKELVSTIMEOLKSKTLFSLMQYSEEF 196

QY 181 RHFTFKFQNNPNRSLIKPTTQLGRTHATGLRKVVRELFTNIINGARKNAFKILFL 240
Db 197 RHFTFKFQNNPNRSLIKPTTQLGRTHATGLRKVVRELFTNIINGARKNAFKILFL 256

QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 300
Db 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVLGFGDAFRSEKSRQELNTVASKPRDHVQAN 316

QY 301 NPEALKTVONQUREKIFAIEGTQTGSSSFHEHMSQEGFSAITSNGLPLSTVGSYDNAG 360
Db 317 NPEALKTVONQUREKIFAIEGTQTGSSSFHEHMSQEGFSAITSNGLPLSTVGSYDNAG 376

QY 361 GVFLYTSKESKTFINMTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKESKTFINMTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVVCP 496

QY 481 PRGQARWQCDVLYEQQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVILF 540
Db 497 PRGQARWQCDVLYEQQCPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVILF 556

QY 541 HGTSGSGISPSHSORITAGSKLSPLQYFGQSLSGGQDLTMDGLVLDITVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSORITAGSKLSPLQYFGQSLSGGQDLTMDGLVLDITVGAQGHVLLRSQ 616

QY 601 PVLVRKAIAMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 660
Db 617 PVLVRKAIAMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 676

QY 661 YDLALDSGRPHSAVENETKNSRTRQTVLGLTQCTETLKLQLPNCIEDPVSPIVLRNLF 720
Db 677 YDLALDSGRPHSAVENETKNSRTRQTVLGLTQCTETLKLQLPNCIEDPVSPIVLRNLF 736
```

```
197 RHFTFKFQNNPNSRLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDGAFRSEKSRQELMTNVASKPPRDHVFOAN 300
257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELMTIASKPPRDHVFOVN 316
301 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLSTVGSYDWAG 360
317 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLSTVGSYDWAG 376
361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 480
437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 496
481 PRGORARQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
497 PRGORARQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 600
557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 616
601 PVLVRKALMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
617 PVLVRKALMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
721 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITFFSMSLDCLVVG 780
737 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITFFSMSLDCLVVG 796
781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACESASSTEV 856
841 SGALKSTSCSINHPIFPENSEVTNIIPDVDSKASLGNKLLKANVTSENMMPTNKTEF 900
857 SGALKSTSCSINHPIFPENSEVTNIIPDVDSKASLGNKLLKANVTSENMMPTNKTEF 916
901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQQVSNLQORSLSPISLVFLVPV 960
917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHQQVSNLQORSLSPISLVFLVPV 976
961 RLNCTVWDPRQVTFSENLSSTCHTKRLPSHSDFLAELRKAPVNCIAVCQRIQCDIP 1020
977 RLNCTVWDPRQVTFSENLSSTCHTKRLPSHSDFLAELRKAPVNCIAVCQRIQCDIP 1036
1021 FFGIQEEFNATLKGNSLDFDWIKTSHNHLITVSTAELFNDVSFTLLPGQAFVRSQDET 1080
1037 FFGIQEEFNATLKGNSLDFDWIKTSHNHLITVSTAELFNDVSFTLLPGQAFVRSQDET 1096
1081 KVERPEFNPPLPLIVGSSVGGLLLLALITAALYKLGFFKQYKDMMEGGPPGAEPPQ 1137
1097 KVERPEFNPPLPLIVGSSVGGLLLLALITAALYKLGFFKQYKDMMEGGPPGAEPPQ 1153
```

## RESULT 6

US-10-144-259-30

; Sequence 30, Application US/10144259

; Publication No. US20030109691A1

; GENERAL INFORMATION:

; APPLICANT: Ainaout, M. Amin

; APPLICANT: Li, Rui

; APPLICANT: Xiong, Jian-Ping

; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF

```
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30
```

```
Query Match 99.3%; Score 5839; DB 14; Length 1153;
Best Local Similarity 98.9%; Pred. No. 0;
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQRGSLVQCDSYSGSCPEI 60
DB 17 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQRGSLVQCDSYSGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTSENTRYVVKLCFLGSLNRQQOK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTSENTRYVVKLCFLGSLNRQQOK 136
QY 121 FPEALRGCPQSDSIAFLIDGSGSIIIPHDFRMKELVSTIMEQLKSKTFLSLMOYSBEF 180
DB 137 FPEALRGCPQSDSIAFLIDGSGSIIIPHDFRMKEFVSTVMEQLKSKTFLSLMOYSBEF 196
QY 181 RIHTTFKEFQNNPNSRLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 240
DB 197 RIHTTFKEFQNNPNSRLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVLGFDGAFRSEKSRQELMTNVASKPPRDHVFOAN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELMTNVASKPPRDHVFOAN 316
QY 301 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLSTVGSYDWAG 360
DB 317 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEOTRGQGVSCPL 496
QY 481 PRGORARQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
DB 497 PRGORARQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHQSRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLVGAQGHVLLRSQ 616
QY 601 PVLVRKALMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
DB 617 PVLVRKALMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720
DB 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736
QY 721 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITFFSMSLDCLVVG 780
DB 737 SLVGTPLSAFNGLRPVLAEDAQRFLTALFPPEKNCNDNICQDDLSITFFSMSLDCLVVG 796
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWRLACESASSTEV 840
```

Db 797 GPRENVTVTVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNSORSWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYINFTASENTRVMQHOYQVSNLQORSLSPLSLVPLVPV 960  
Db 917 QLELPVKYAVYVMTSHGVSTKYINFTASENTRVMQHOYQVSNLQORSLSPLSLVPLVPV 976  
QY 961 RLNQTVIWDPRQVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQCDDIP 1020  
Db 977 RLNQTVIWDPRQVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQCDDIP 1036  
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFVLLPQGGAFVRSQTET 1080  
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFVLLPQGGAFVRSQTET 1096  
QY 1081 KVEPFEVNPPLIYGVSSVGGILLALITAAALYKLGFFKRYKDMSEGGPPGABPQ 1137  
Db 1097 KVEPFEVNPPLIYGVSSVGGILLALITAAALYKLGFFKRYKDMSEGGPPGABPQ 1153

RESULT 7  
US-10-207-655-176  
; Sequence 176, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069.401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 176  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-207-655-176

Query Match 99.3%; Score 5839; DB 14; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAANQSGLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAANQSGLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQCSENTYKGLCFLFGSNLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQCSENTYKGLCFLFGSNLRQOPQK 136  
QY 121 FPEALRGCCQEDSDIAFLIDGSGIIPHPRRMKELVSTIMEOLKSKTLFSLMQYSEEF 180  
Db 137 FPEALRGCCQEDSDIAFLIDGSGIIPHPRRMKELVSTIMEOLKSKTLFSLMQYSEEF 196  
QY 181 RIHFTKPFQNNPNRSLKIPITOLLGRTHATGLARKVRELFTNITNGARKNAFKILFL 240  
Db 197 RIHFTKPFQNNPNRSLKIPITOLLGRTHATGLARKVRELFTNITNGARKNAFKILVI 256  
QY 241 TDGEKGDPLGYEDVITPELDREGVIRYVGLFGDAFSEKSRQELNTVASKPPEDHVFQAN 300  
Db 257 TDGEKGDPLGYEDVITPELDREGVIRYVGLFGDAFSEKSRQELNTVASKPPEDHVFQAN 316  
QY 301 NFPAKTVQNLREKIFAEGTQGTGSSSSPEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360  
Db 317 NFPAKTVQNLREKIFAEGTQGTGSSSSPEHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSVLGLAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSVLGLAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYQTRGGQSVCP 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYQTRGGQSVCP 496  
QY 481 PRGQRRQCDVAVLYGEOQPGWRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAAYLF 540  
Db 497 PRGQRRQCDVAVLYGEOQPGWRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAAYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGSLSGQDITMDGLVDTLTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGSLSGQDITMDGLVDTLTVGAQGHVLLRSQ 616  
QY 601 PVLRYKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRRLRREGQIQSVVT 660  
Db 617 PVLRYKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRRLRREGQIQSVVT 676  
QY 661 YDLALDSGRPHSRAVFNETKNSRRQTVLGTQTCETLKLQLPNCIEDPVPSPVILRNLF 720  
Db 677 YDLALDSGRPHSRAVFNETKNSRRQTVLGTQTCETLKLQLPNCIEDPVPSPVILRNLF 736  
QY 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALPPFEKNCNDNICODDLSITFSEMSLDCLVVG 780  
Db 737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALPPFEKNCNDNICODDLSITFSEMSLDCLVVG 796  
QY 781 GPRENVTVTVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNSORSWRLACESASSTEV 840  
Db 797 GPRENVTVTVRNDEGDSYRTQVTFPPFLDLISYRKVSTLQNSORSWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYINFTASENTRVMQHOYQVSNLQORSLSPLSLVPLVPV 960  
Db 917 QLELPVKYAVYVMTSHGVSTKYINFTASENTRVMQHOYQVSNLQORSLSPLSLVPLVPV 976  
QY 961 RLNQTVIWDPRQVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQCDDIP 1020  
Db 977 RLNQTVIWDPRQVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQCDDIP 1036  
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFVLLPQGGAFVRSQTET 1080  
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLLIVSTAEILFNDVSFVLLPQGGAFVRSQTET 1096  
QY 1081 KVEPFEVNPPLIYGVSSVGGILLALITAAALYKLGFFKRYKDMSEGGPPGABPQ 1137  
Db 1097 KVEPFEVNPPLIYGVSSVGGILLALITAAALYKLGFFKRYKDMSEGGPPGABPQ 1153

## RESULT 8

US-09-902-481A-5  
; Sequence 5, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic

US-09-902-481A-5

Query Match 99.3%; Score 5836; DB 10; Length 1137;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVQCDYSTGSCPEI 60  
DB 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVQCDYSTGSCPEI 60

QY 61 RLQVPVEAVNMVGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFGLFSGNLRQOPQK 120  
DB 61 RLQVPVEAVNMVGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFGLFSGNLRQOPQK 120

QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEP 180  
DB 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEP 180

QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240  
DB 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVGLFGDAPRSEKSRQELNVTASKPPRDHVFOAN 300  
DB 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNVTASKPPRDHVFOIN 300

QY 301 NFEALKTQVONLRKIPFAIEGTOTGSSSSFEHMSOEGFSAITNSGPLLSTVGSYDWAG 360  
DB 301 NFEALKTQVONLRKIPFAIEGTOTGSSSSFEHMSOEGFSAITNSGPLLSTVGSYDWAG 360

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420  
DB 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480  
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480

QY 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVLGVNDGDKLTDVAIGAPGEDNKGAVYLF 540  
DB 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVLGVNDGDKLTDVAIGAPGEDNKGAVYLF 540

QY 541 HGTSGSGISPSHSQRIAGSKLSPLQLVFGQSLSGQDLTMDGLVDLTVGAGCHVLLLRSQ 600  
DB 541 HGTSGSGISPSHSQRIAGSKLSPLQLVFGQSLSGQDLTMDGLVDLTVGAGCHVLLLRSQ 600

QY 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 601 PVLRVKAIMFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660

QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCTEDPSPVILRLNF 720  
DB 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQPNCTEDPSPVILRLNF 720

QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITTFSPMSLDCLVVG 780  
DB 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITTFSPMSLDCLVVG 780

QY 781 GPREFNVTVRNDGEDSYQTQVTFPPPLDLSYKVKSTLQNRQSRQWRLACESASSTEV 840  
DB 781 GPREFNVTVRNDGEDSYQTQVTFPPPLDLSYKVKSTLQNRQSRQWRLACESASSTEV 840

QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
DB 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900

QY 901 QLELPVYAVYVMVTSHGVSSTKYLNFNTASNTSRVMQHVQVSNLQORSPLISLVFLVPV 960  
DB 901 QLELPVYAVYVMVTSHGVSSTKYLNFNTASNTSRVMQHVQVSNLQORSPLISLVFLVPV 960

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
DB 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020

QY 1021 PFGIOEBFNATLKGNSLSPDWIKTSHNHLILVSTAEIILFNDVSFTLLPQCGAFVRSQTET 1080  
DB 1021 PFGIOEBFNATLKGNSLSPDWIKTSHNHLILVSTAEIILFNDVSFTLLPQCGAFVRSQTET 1080

QY 1081 KVEPFEVFNPLPLIVGSSVGGLLILALITALYKLGFEKROYKDMWSEGGPPGABPO 1137  
DB 1081 KVEPFEVFNPLPLIVGSSVGGLLILALITALYKLGFEKROYKDMWSEGGPPGABPO 1137

## RESULT 9

US-09-902-481A-6  
; Sequence 6 Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902.481A  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 6  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-6

Query Match 99.2%; Score 5832; DB 10; Length 1137;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVQCDYSTGSCPEI 60  
DB 1 FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPOEIVAAQNRGSLVQCDYSTGSCPEI 60

QY 61 RLQVPVEAVNMVGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFGLFSGNLRQOPQK 120  
DB 61 RLQVPVEAVNMVGLSLAATTSPPOLLACGPTVHQTCSNTYVVKGLCFGLFSGNLRQOPQK 120

QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEP 180  
DB 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTFLSLMOYSEEP 180

QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240  
DB 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240

QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVGLFGDAPRSEKSRQELNVTASKPPRDHVFOAN 300  
DB 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNVTASKPPRDHVFOVN 300

QY 301 NFEALKTQVONLRKIPFAIEGTOTGSSSSFEHMSOEGFSAITNSGPLLSTVGSYDWAG 360  
DB 301 NFEALKTQVONLRKIPFAIEGTOTGSSSSFEHMSOEGFSAITNSGPLLSTVGSYDWAG 360

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420  
DB 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480  
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480

QY 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVLGVNDGDKLTDVAIGAPGEDNKGAVYLF 540  
DB 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVLGVNDGDKLTDVAIGAPGEDNKGAVYLF 540



```
Db 481 PRGQARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
QY 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGOSLGGQDLTMGCLVDLTGVAQGHVLLRSQ 600
Db 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGOSLGGQDLTMGCLVDLTGVAQGHVLLRSQ 600
QY 601 PVLVRKAMERNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
Db 601 PVLVRKAMERNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVPVLRNF 720
Db 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVPVLRNF 720
QY 721 SLVGTPLSAFGLRPLVAEDAQRLETFALPPFKNGCNDNICODDLSITFSFMSLCLVVG 780
Db 721 SLVGTPLSAFGLRPLVAEDAQRLETFALPPFKNGCNDNICODDLSITFSFMSLCLVVG 780
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLVSRKVSTLQORSQSWRLACESASSTEV 840
Db 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLVSRKVSTLQORSQSWRLACESASSTEV 840
QY 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFLVPV 960
Db 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFLVPV 960
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
Db 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
QY 1021 PFGIQEEFNATLKGNLSPDWYIKTSHNLLIIVSTAEILFNDSVFTLLPQCGAFVRSQTEP 1080
Db 1021 PFGIQEEFNATLKGNLSPDWYIKTSHNLLIIVSTAEILFNDSVFTLLPQCGAFVRSQTEP 1080
QY 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGGPPGAEPQ 1137
Db 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGGPPGAEPQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shinooka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-945-265-4

Query Match 99.1%; Score 5823.5; DB 9; Length 1152;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 1123; Conservative 8; Mismatches 5; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAVNRGSLYQCDYSTGSCPEI 76
```

```
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCFGLFSLNRQOQPK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCFGLFSLNRQOQPK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKELVSTIMEQLKKSKTLFSLMYQSEBF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQLKKSKTLFSLMYQSEBF 196
QY 181 RIHFTFKFQNNPNRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPELDRGVIRYVLFDFDAPRSEKSRQELNATVASKPRDHVFOAN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVGVGDAPRSEKSRQELNATASKPRDHVFOAN 316
QY 301 NFEALKTQVONQUREKIFAIEGTQTGSSSPHEMSQEGFSAAITNGPILLSITVGYSDWAG 360
Db 317 NFEALKTQVONQUREKIFAIEGTQTGSSSPHEMSQEGFSAAITNGPILLSITVGYSDWAG 376
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPYQHIGLVAMER 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPYQHIGLVAMER 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSVCP 496
QY 481 PRGQARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRG-RARWQCDVAVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
QY 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGOSLGGQDLTMGCLVDLTGVAQGHVLLRSQ 600
Db 556 HGTSGSGISPHSQRIAGSKLSPRLQYFGOSLGGQDLTMGCLVDLTGVAQGHVLLRSQ 615
QY 601 PVLVRKAMERNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 660
Db 616 PVLVRKAMERNPREVARNVFCNDQVVKGEAGVRVCLHVOKSTRDLREGQIQSVVT 675
QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVPVLRNF 720
Db 676 YDLALDSGRPHSRVAFNETKSTRQTQVLGTQTCETLKLQLPNCIEDPVPVLRNF 735
QY 721 SLVGTPLSAFGLRPLVAEDAQRLETFALPPFKNGCNDNICODDLSITFSFMSLCLVVG 780
Db 736 SLVGTPLSAFGLRPLVAEDAQRLETFALPPFKNGCNDNICODDLSITFSFMSLCLVVG 795
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLVSRKVSTLQORSQSWRLACESASSTEV 840
Db 796 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLVSRKVSTLQORSQSWRLACESASSTEV 855
QY 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 856 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFLVPV 960
Db 916 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSPLSLVFLVPV 975
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1020
Db 976 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDDIP 1035
QY 1021 PFGIQEEFNATLKGNLSPDWYIKTSHNLLIIVSTAEILFNDSVFTLLPQCGAFVRSQTEP 1080
Db 1036 PFGIQEEFNATLKGNLSPDWYIKTSHNLLIIVSTAEILFNDSVFTLLPQCGAFVRSQTEP 1095
QY 1081 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGGPPGAEPQ 1137
Db 1096 KVEPPEVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKRYQKDMMSGGPPGAEPQ 1152
```



```
RESULT 11
US-10-116-275-204
; Sequence 204, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; TITLE OF INVENTION: Compositions Targeting Peyer's Patches and M Cell Receptors
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116, 275
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 204
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-116-275-204

Query Match      58.6%; Score 3446; DB 14; Length 1163;
Best Local Similarity 60.5%; Pred. No. 1.2e-284;
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGQSVVQLOGSRVVGAPQEIIVANQORSLYQCDYSTGSCPEI 60
Db 20 FNLDTTELTAFRVDSAGFSDSVVQYANVWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKGLCPFGSNLRQOPQK 120
Db 80 GLQVPPAVNMSLGLSLAATTSPOLLACGPTVHHECGRNMYLTGLCFLGPT--QLTOR 137

QY 121 FPEALRGCPQEDSIAFLIDGSGSIIIPHDFRMKELVSTIMEQLKSKTLFSLMOYSEF 180
Db 138 LPVSRQCEPRQEQDIFLIDGSGSISSENFMVFAVVISQRPSTQSLMQFSNKF 197

QY 181 RIHTFQFQNNPRSLIKIPITOLLGTHATGLRKVVRELFINITNGARKNAFKILFL 240
Db 198 QTHFTFEFRSTNPLSLLASVHQLQGYTATATQNVVHLEFASVAGRDATKILIVI 257

QY 241 TDGKFGDPLGYEDVIELDREGVIRYVGLFGDAFRSEKSOELNTYASKPRDHVQAN 300
Db 258 TDGKKGDSLDYKDVIPWADAGIIRVAGVGLAFQPNRNNSWKLNDIASKPSQEHIFKVE 317

QY 301 NFEALKTQVQNLREKIFAIEGTOTGSSSFEHMSQEGFSAATISNGPELLSTVGSYDWAG 360
Db 318 DFDALKDIQNLKEKIFAIEGTETTSSTSSFELEMAQEGFSAVFTPDGPFVLAGVGSFTWSG 377

QY 361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGVAALILNRRVQSLVLAGAPRYOHIGLVAMFR 420
Db 378 GAFLYPPNMGPTFINMSQENVMDSDYLGYSTELALWKGVQSLVLAGAPRYOHTKAVIFT 437

QY 421 QNTGMESNANVKTQIGVFGASLCSVDVDSNGSTDVILIGAPHYYEOTRGQVSVCP 480
Db 438 QVSEQWKMKEVGTQIGVFGASLCSVDVDTGSDTLVILIGAPHYYEOTRGQVSVCP 497

QY 481 PRQARWQCDVILYGRQGFQWGRFGAALTVLGVNDGDKLTDVAIGAPGEDNRRGAYVLF 540
Db 498 PRGWR-RWMCDAVLYGQGHQWGRFGAALTVLGVNDGDKLTDVVGAPGEENRGAYVLF 556

QY 541 HGTSGSISPSHSORVAGSLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVILLRSQ 600
Db 557 HGVLPISPSHSORVAGSLSPRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLLRTR 616

QY 601 PVLVRKAIMENPREVARNFECDNDVYKGEAGEVRVCHVQKSTRDRIRBEGIQSVVT 660
Db 617 PVLVGVSMQFIPAEIPRSAFEQEVVSEQTLVQSNICLYIDKRKNLLGSRDQSSVT 676
```

```
QY 661 YDLALDSGRPHSRVFNETHKSTRBQTQVGLGTQTCETLKLQIPNCIEHPVSPVILRNLF 720
Db 677 LDALDPGRLSPRATFOETKNRSLSRVRVGLKAHCENFNLLPSCVEDSVPTILRNLF 736

QY 721 SLVGTPLSAFAGNLRPVLAEDAQRFLTALFPFPEKNCNNDNICODDLSITSEFSLDCLVVG 780
Db 737 TLVGRPLLAFRNLRPMLAALAQRYFTASLPFEKNCADHICQDNIGISFSPGLKSLVVG 796

QY 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYKRVSTLQNRQSRQSRWLACESASSTEV 840
Db 797 SNLELNAEVMVMWNGEDSYGTTITFTHPAGLSYRYVAEGQKQGLRSLHLTCDAPVG-- 854

QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMMRINKTEF 900
Db 855 SQGTWSTSCRINHLLIFRGGAQITFLATFDVSPKAVLGDRLLTANVSENNTPTSKTTF 914

QY 901 QLELPVYAVYVVTSHGVSTKYLNFTAS-ENTSRVMOHOYQVSNLQORSLSPISELVLP 959
Db 915 QLELPVYAVYVTVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSINFVWP 974

QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSHDSDFLAELKAPVNVNCSIAVCQIQCDI 1019
Db 975 VELNQEAVMMDVEVSHQPONPSLRCSSEKIAAPPASDFLAHIQKNPVLDCSIAGCLFRCDV 1034

QY 1020 PFFGIQEEFNATLKGNSLSDFWIKTSHNHLIIVSTAELLFNDSVFTLLPGQAFVRSOTE 1079
Db 1035 PSFVQBELDFTLKGNSLFGWVRQILOKKVSVSWVAEITFTDTSVYSQJPGQEAFFMRAOTT 1094

QY 1080 TKVPEFVEPNPLIIVGSSVGGLLLALITAAALKGLGFFKRYQKDMSE 1128
Db 1095 TVLEKYKHNPPTLIIVSSIGGLLLALITAVLYKVGFKKRYKEMEE 1143
```

```
RESULT 12
US-09-350-259-4
; Sequence 4, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350, 259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193, 043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173, 497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286, 889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362, 652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943, 363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-259-4
```

```
Query Match      58.2%; Score 3423; DB 9; Length 1163;
Best Local Similarity 60.2%; Pred. No. 1.1e-282;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGQSVVQLOGSRVVGAPQEIIVANQORSLYQCDYSTGSCPEI 60
Db 20 FNLDTTELTAFRVDSAGFSDSVVQYANVWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKGLCFLGSLNRQOPQK 120
Db 80 GLQVPPAVNMSLGLSLAATTSPOLLACGPTVHHECGRNMYLTGLCFLGPT--QLTOR 137
```

```
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTSLSLMOYSEEF 180
Db 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFAATMNFVRAVISQFORPSTQFSLMQFSNKF 197
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGURKVVRELFNTNGARKNAFKILFL 240
Db 198 QTHFTFEFRRTSNPLSLASVHQLGFTYTATLQNVVHRLFHASYGARRDAIKILIVI 257
QY 241 TDGERFGDPLGVEDVPELDRBGVIRYVLGDFGDFRSEKSRQELNTVASKPPDRHVFQAN 300
Db 258 TDGKKEGSDLDYKDVIPMDAAGIIRYALGVGLAFQNRNSWKLNDIASKPSQEHFKVE 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTSSEFHEMSQEGFSAITNGPLLSVGSYDWAG 360
Db 318 DFDALKDIONQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTDPGVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINNSQENVDMRDSYLGSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQVSVCP 480
Db 438 QVSRQWEMKAEVIGTOIGSYFGASLCSVDVDTGSDTDLVIGAPHYYEOTRGQVSVCP 497
QY 481 PRQARWOCDAVLYGEOQGPWRFGAALTVLGVNGDKLTDVAIGAPGEENRGAVYLF 540
Db 498 PRGWR-RWNCDAVLYGEOQHPWGRFGAALTVLGVNGDKLTDVVI GAPGEENRGAVYLF 556
QY 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSLGGODLTMDGLVDTLVCAGHVLRLRSQ 600
Db 557 HGVLGPSISPSHSORITAGSKLSPRLQYFGOALSQQDGLTDGLVLDVAVGARGVLLLR 616
QY 601 PVLVRKAIMFNPREVARNVFECDQVWVGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLWGVGSMQFIPAEIPRSAFECEQVWSEQTILVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSAVNFNETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSFIVRLNF 720
```

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

```
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: NO. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-09-891-943-4
```

```
Query Match 58.2%; Score 3423; DB 10; Length 1163;
Best Local Similarity 60.2%; Pred. No. 1.1e-282;
Matches 680; Conservative 149; Mismatches 294; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARGGQSVQVQCGSRVVVGAPQEIIVANQORSLYQCDYSTGSCPEI 60
Db 20 FNLDTTELTAFRVDSAGFGSDSVVQYANSWVVVGAFQKIIAANQIGSLYQCGYSTGACPEI 79
QY 61 RLQVPEAVNMSLGLSLAATTPSPOLLACGPVTHQTSNTYVKGLCFFGSLNLRQPOPK 120
Db 80 GLQVPEAVNMSLGLSLAATTPSPOLLACGPVTHQTSNTYVKGLCFFGSLNLRQPOPK 137
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEQLKSKTSLSLMOYSEEF 180
Db 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFAATMNFVRAVISQFORPSTQFSLMQFSNKF 197
QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGURKVVRELFNTNGARKNAFKILFL 240
Db 198 QTHFTFEFRRTSNPLSLASVHQLGFTYTATLQNVVHRLFHASYGARRDAIKILIVI 257
QY 241 TDGERFGDPLGVEDVPELDRBGVIRYVLGDFGDFRSEKSRQELNTVASKPPDRHVFQAN 300
Db 258 TDGKKEGSDLDYKDVIPMDAAGIIRYALGVGLAFQNRNSWKLNDIASKPSQEHFKVE 317
QY 301 NFEALKTQVQNLREKIFAIEGTQTSSEFHEMSQEGFSAITNGPLLSVGSYDWAG 360
Db 318 DFDALKDIONQLKEKIFAIEGTETISSSSFELEMAQEGFSAVFTDPGVLGAVGFTWSG 377
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINNSQENVDMRDSYLGSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQVSVCP 480
Db 438 QVSRQWEMKAEVIGTOIGSYFGASLCSVDVDTGSDTDLVIGAPHYYEOTRGQVSVCP 497
QY 481 PRQARWOCDAVLYGEOQGPWRFGAALTVLGVNGDKLTDVAIGAPGEENRGAVYLF 540
Db 498 PRGWR-RWNCDAVLYGEOQHPWGRFGAALTVLGVNGDKLTDVVI GAPGEENRGAVYLF 556
QY 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGOSLGGODLTMDGLVDTLVCAGHVLRLRSQ 600
Db 557 HGVLGPSISPSHSORITAGSKLSPRLQYFGOALSQQDGLTDGLVLDVAVGARGVLLLR 616
QY 601 PVLVRKAIMFNPREVARNVFECDQVWVGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLWGVGSMQFIPAEIPRSAFECEQVWSEQTILVQSNICLYIDKRSKNLLGSRDLQSSVT 676
QY 661 YDLALDSGRPHSAVNFNETKNSTRQTVLGLTQTCETLKLQPLNCIEDPVSFIVRLNF 720
```

Db 677 LDALAPGRSLSPRAIFQETKNSRSLRVGLKAHCENFNLLPSCVEDSVIPIILRLNF 736  
Qy 721 SLVGTPLSAGNLRVLAEDAQRLFTALPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
Db 737 TLVGGKLLAFRLRPMALAAQRYFTASLPFEKNGADHI CODNIGISFSFPLKSLVVG 796  
Qy 781 GPREFNVTVVRNDCGEDSYRTQVTFPFLDLISYKVVSTLQNRORSWRLACESASSTEV 840  
Db 797 SNLELNAEVMWVNDGSDSYGTTITESHAGLSYRVVARGQKQQLRSLLHLC--CSAPVG 854  
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVSKASLGNKLLKANVTSENMPRTNKTEF 900  
Db 855 SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENNIPRTSKTIF 914  
Qy 901 QLELPVKYAVVMVYTSHGVSNTKYLNF--ENTSRVMOHQVSNLQORSPISLVFLVP 959  
Db 915 QLELPVKYAVVMVYTSHGVSNTKYLNFSESEKESHVAMHYQVNLGQDLFVSNFVWP 974  
Qy 960 VRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORICDI 1019  
Db 975 VELNQEAVMWVESHQPNPSLRCSEKXIAPPASDFLAHQNPVLDCSIAGLRFRCDV 1034  
Qy 1020 PFGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079  
Db 1035 PSFSVQEBDLFTLKNLSFGWVRQILQKKSWSVSAEIIIFDTSVYSQPLGQEFMRAQTI 1094  
Qy 1080 TKVPEFVNPPLVTCVSSVGLLALLALITAAALYKLGFFKQYKDMXSE 1128  
Db 1095 TVLEKIKVHNPIPLTVGSSIGGLLLALLITAVLYKVGFFKQYKEMMBE 1143

## RESULT 14

US-09-350-259-2  
; Sequence 2, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20020062008A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1161  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-2

Query Match 57.6%; Score 3388; DB 9; Length 1161;  
Best Local Similarity 58.9%; Pred. No. 1,le-279;  
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;  
Qy 1 FNLDTENAMTFQENARFGGSVWOLQGSRRVVGAPQFIVAANQRSGLYQCDDYSTGSCPEI 60  
Db 17 FNLDVEEPTIFQEDAGFGGSVWVFGSRLVVGAPLEVAANQTRGLYDCAAAATGMCQPI 76  
Qy 61 RLQVPEAVNMSGLSLAATTSPPOLACGPTHQTCSENTYUKGLCFGLFPLNRQOPQK 120  
Db 77 PLHIRPEAVNMSGLTLAATNSGRLLACGPTHVRVCGENSYSKSGSCLLLGSRW-ELIQT 135

## RESULT 15

US-09-891-943-2  
; Sequence 2, Application US/09891943  
; Publication No. US2003007728A1

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSII PHDFRMKELVSTIMEOLKSKTILFSIMOYSBEF 180  
Db 136 VPDATPBCPHQEMDVLFLIDGSGSDQDNFQMGFVQAVMGQFEGTDTTLFALMOYSNLL 195  
Qy 181 RIHFTTFKEFONNPNRSLIKPITOLLGRTHATGLRKVVRELFNITNGARKNAPKILFLL 240  
Db 196 KIHFTTFOFRISPSQOSLVDPIVLKGLTFTATGLITVVTQLFHFKNGARKSAKKILV 255  
Qy 241 TDGSKFGDPLGYEDVIPELDREGVIRVVLGFGDAPRSEKSRQELNVTASKPRDRHFQAN 300  
Db 256 TDGQYKIDPLEYSDVIPAQKAGIIRVAIGVGHAFQGTARQELNVTISSAPPQDHVFXVD 315  
Qy 301 NFEALKTVQNLREKIFALECTQTGSSSSPHEMSOEGFSAAITNSGPLLSTVGSYDWAG 360  
Db 316 NFAALGSIQOLQKBIYAVEGTQSRASSSFOHMSOEGFSTALTMDGFLGAVGSFWSG 375  
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDMDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420  
Db 376 GAFLYPPNMSPTFINMSQENVDMEDSVLGYSTELALMKGVQNLVLGAPRYOHTGKAVFT 435  
Qy 421 QNTGWSNSANVKGTQIGAYGASLCSVDVDSNGSTDVLVLIGAPHYYEQTRGGQVSCPL 480  
Db 436 QVSRQWRKKAEBVTGTGYSYFGASLCSVDVDSGSTDLLILGAPHYYEQTRGGQVSCPL 495  
Qy 481 PRGORAEWQCDVILYGGQGPWGFAGALTVLGVDNGDKLTDVAIGAPGEDNKGAVLYF 540  
Db 496 PRGORVQWQCDVILYGGQGPWGFAGALTVLGVDNEDKLDVAIGAPGEQNGAVLYF 555  
Qy 541 HGTSGSGISPSHSORIASGSKLSPRLQYFGOSLGGQDLTMDGLVDTLTVGAOGHVLLLRSQ 600  
Db 556 HGASESGISPSHSORIASQSLSPRLQYFGQALSGQDLTQDGLMDLAVARGQVLLLRSL 615  
Qy 601 PVLVKVKAINEFNPREVARNVPCNDQVVKGEAGEVVRVCLHVQKSTRDRRREGQIQSVVT 660  
Db 616 PVLKGVAMRFSPEVAVAKVYRCWEEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673  
Qy 661 YDLALDSGRPHSRAVENETKNSTRQVQLGTQTCETLKLQENCLDPSVTLRLNPF 720  
Db 674 FDLALDPRLTSTRAIFNETKNPTLTRKTLGLGHCTELKLLPDCVEDVVSPIILHLNF 733  
Qy 721 SLVGTPLSFAFNLRPVLAEDAQRFLTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780  
Db 734 SLVREPIPSQNLRPVLAVGSQDLFTASLPFEKNCQDGLCEGDLGVTLSFGLOTITVG 793  
Qy 781 GPREFNVTVVRNDCGEDSYRTQVTFPFLDLISYKVVSTLQNRORSWRLACESASSTEV 840  
Db 794 SSLELNVITVMNAGEDSYGTWVSLYYPAGLSHRRVSGAQKPHQSALRLACETV-PTED 852  
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVSKASLGNKLLKANVTSENMPRTNKTEF 900  
Db 853 EG-LRSCRCVSNHPIFHEGNGTFTVTFDVSYKATLGDRLMLRASSENKASSSKATF 911  
Qy 901 QLELPVKYAVVMVYTSHGVSNTKYLNF--TASENTSRVMOHQVSNLQORSPISLVFLVP 959  
Db 912 QLELPVKYAVVMVYTSHGVSNTKYLNF--TASENTSRVMOHQVSNLQORSPISLVFLVP 971  
Qy 960 VRLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORICDI 1019  
Db 972 VLLNGVAVMDVWMEAPSQSL--PCVSEKPPQHSDFLTQISRSEMLDCSIADCLQFPCDV 1029  
Qy 1020 PFGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSQTE 1079  
Db 1030 PSFSVQEBDLFTLKNLSFGWVRQILQKKSWSVSAEIIIFDTSVYSQPLGQEFMRAQME 1089  
Qy 1080 TKVPEFVNPPLVTVGSSVGLLALLALITAAALYKLGFFKQYKDMXSE 1128  
Db 1090 MVLEEDVYNAIPIINGSSVGLLALLALITATLYKLGFFKRYKHEKMLED 1138

GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20030077278A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/891,943  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1161  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-891-943-2

Query Match 57.6%; Score 3388; DB 10; Length 1161;  
Best Local Similarity 58.9%; Pred. No. 1.1e-279;  
Matches 665; Conservative 168; Mismatches 288; Indels 8; Gaps 6;

QY 1 FNLDTENAMTFOENARGFGQSVVLQGGSRVVVGAPOEIVAAANQORSLYOCYSTGSCPEI 60  
DB 17 FNLDEEPTIFQEDAGGFGQSVVQGGSRVVVGAPOEIVAAANQORSLYOCYSTGSCPEI 76

QY 61 RLOVPEAVNMSLGLSLAATTSPPOLLACGPTVHOCSTENTYVKGICFLFGSLNQPOK 120  
DB 77 PCHIRPEAVNMSLGLSLAATTSPPOLLACGPTVHOCSTENTYVKGICFLFGSLNQPOK 135

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELIVSTIMBQLKKKTLFSLMOYSEEF 180  
DB 136 VPDATPECPHOEMDIVFLIDGSGSIDQDNFNQMGFVQAVMGQFEGTDTLFAIMOYSNLL 195

QY 181 RHFTFKFQNNPNRSLIKPTOLLGRTHATGLKVVRELPNITNGARKNAFLIFLL 240  
DB 196 KIHFTFTQPTSPSQSLVDPIVLQKGLTFTATGILTVTQLFHHKNGARKAKKILIVI 255

QY 241 TDGKFGDPLGVEDVPELDRSGVIRYVLGFGDFPESEKSRORLNTVASKPRDHFVOAN 300  
DB 256 TDGQKYKPLEYSDVPIQAEKAGIIRYALGVGHAFQPTARQELNTISSAPQDHFVKVD 315

QY 301 NFEALKTQNLREKIFALEGTQTSSESSFEHMSQEGFSAATSNGLLSTVGSYDWAG 360  
DB 316 NFAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMGDLFLGAVGSPSWG 375

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAIILRNQSVLVLGAPRYOHIGLVAMFR 420  
DB 376 GAFLYPPNMSPTFINNSQENVMDRSDYLGSTELALWKGVQNLVLGAPRYOHTGKAVIF 435

QY 421 QNTGMWESNANVKGTOIGAVFGASLCSDVDVNSGTDLVLIGAPHYVEOTRGQVSVCP 480  
DB 436 QVSRQWRKAEVGTGQIGSYFGASLCSDVDVNSGTDLVLIGAPHYVEOTRGQVSVCP 495

QY 481 PRGQARWOCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 496 PRGQVQWCCDAVLRGEQGHMGRFGAALTVLGDVNNEDKLDVAIGAPGEENRGAVYLF 555

QY 541 HGTSGSGISPSHSQRIAGSLSPRLQYFGOSLGGQDLTMDGLVDLTGVAQGHVILLRSQ 600  
DB 556 HGASESGISPSHSQRIAGSLSPRLQYFGOSLGGQDLTMDGLVDLTGVAQGHVILLRSQ 615

QY 601 PVLVRKAIMFENPREAVNVECNQVKGKEGVRVCLVQVKSTRDLRREGQIQSVVT 660  
DB 616 PVLKVGVMFSPVEVAKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

QY 661 YDLALDGRPHSRVAVNETKNSRTRQTVLGLTQTCTETIKLQLPNCIEDPVSPIVLRNPF 720

Db 674 FDLALDGRHLSRAIFNETKNPTLTRRXTLGLGHCETLKLALLPDCVEDVVVSPILHLNF 733  
QY 721 SLVGTPLSAFCNLRPVLAEADAQRLFTALFPPEKNCNGDNI CODDLSITFSEFMSLDCLVVG 780  
Db 734 SLVREPSPQNLPRPVLAVGSDLFATSLPPEKNCNGDCLCEGLGVTLSFSGLTIVG 793  
QY 781 GPRBFNVTIVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840  
Db 794 SSLELNVITVWNAGEDSYGTIVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852  
QY 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRINKTEF 900  
Db 853 EG-LRSSRCSVNHPIFHEGSGNGTFTVTFDVSYKATLGRMLMRASSENKASSKATF 911  
QY 901 QLELPVKYAVVMVYVTSHGVSSTKYNF-TASENTSRVMOHQVSNLQORSIPISLVFLVP 959  
Db 912 QLELPVKYAVVMVYVTSHGVSSTKYNF-TASENTSRVMOHQVSNLQORSIPISLVFLVP 971  
QY 960 VRLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQDI 1019  
Db 972 VLLNGVAVMDVMEAPSQSL--PCVSEKPKPQHSDFLTQISRSPMLDCSIADCLQRCDV 1029  
QY 1020 PFFGIQEEFNATLKNLSFDDWIKTSHNHLLIVSTAEILFNDSVFTLLPGQGAFFVRSQTE 1079  
Db 1030 PSFSVQEEBLDTLKNLSFGWVRETQKKVLVWSVAEITFDTSVYSQPLPGQGAFFVRSQTE 1089  
QY 1080 TKVEPPEVNPPLPLTVGSSVGLLALLALITAAALYKLGFFKRYKQKMMSE 1128  
Db 1090 MVLDEDEVYNAIPIITGSSVGLLALLALITAAALYKLGFFKRYKQKMMSE 1138

Search completed: November 9, 2004, 12:46:48  
Job time: 102.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds  
(without alignments)  
3278.416 Million cell updates/sec

Title: US-09-902-481B-3

Perfect score: 5879

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMWSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

1: /cgn2\_6/ptodata/1/iaa/5A-COMB.pep.\*

2: /cgn2\_6/ptodata/1/iaa/5B-COMB.pep.\*

3: /cgn2\_6/ptodata/1/iaa/6A-COMB.pep.\*

4: /cgn2\_6/ptodata/1/iaa/6B-COMB.pep.\*

5: /cgn2\_6/ptodata/1/iaa/FACTUS-COMB.pep.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5839	99.3	1153	1 US-08-173-497-3	Sequence 3, Appli
2	5839	99.3	1153	1 US-08-286-889-3	Sequence 3, Appli
3	5839	99.3	1153	1 US-08-485-618-3	Sequence 3, Appli
4	5839	99.3	1153	1 US-08-362-652-3	Sequence 3, Appli
5	5839	99.3	1153	2 US-08-605-672-3	Sequence 3, Appli
6	5839	99.3	1153	2 US-08-482-293A-3	Sequence 3, Appli
7	5839	99.3	1153	2 US-08-943-363-3	Sequence 3, Appli
8	5839	99.3	1153	3 US-09-193-043-3	Sequence 3, Appli
9	5839	99.3	1153	4 US-09-688-307A-3	Sequence 3, Appli
10	5839	99.3	1153	4 US-09-350-259-3	Sequence 3, Appli
11	5808.5	98.8	1152	2 US-08-476-062A-43	Sequence 43, Appl
12	5808.5	98.8	1152	5 PCT-US96-01314-43	Sequence 43, Appl
13	5808.5	98.8	1152	6 5424399-2	Patent No. 5424399
14	3446	58.6	1163	2 US-08-476-062A-44	Sequence 44, Appl
15	3446	58.6	1163	5 PCT-US96-01314-44	Sequence 44, Appl
16	3423	58.2	1163	1 US-08-173-497-4	Sequence 4, Appli
17	3423	58.2	1163	1 US-08-286-889-4	Sequence 4, Appli
18	3423	58.2	1163	1 US-08-485-618-4	Sequence 4, Appli
19	3423	58.2	1163	1 US-08-362-652-4	Sequence 4, Appli
20	3423	58.2	1163	2 US-08-605-672-4	Sequence 4, Appli
21	3423	58.2	1163	2 US-08-482-293A-4	Sequence 4, Appli
22	3423	58.2	1163	2 US-08-943-363-4	Sequence 4, Appli
23	3423	58.2	1163	3 US-09-193-043-4	Sequence 4, Appli
24	3423	58.2	1163	4 US-09-688-307A-4	Sequence 4, Appli
25	3423	58.2	1163	4 US-09-350-259-4	Sequence 4, Appli
26	3388	57.6	1161	1 US-08-173-497-2	Sequence 2, Appli
27	3388	57.6	1161	1 US-08-286-889-2	Sequence 2, Appli

28 3388 57.6 1161 1 US-08-485-618-2 Sequence 2, Appli

29 3388 57.6 1161 1 US-08-362-652-2 Sequence 2, Appli

30 3388 57.6 1161 2 US-08-605-672-2 Sequence 2, Appli

31 3388 57.6 1161 2 US-08-482-293A-2 Sequence 2, Appli

32 3388 57.6 1161 2 US-08-943-363-2 Sequence 2, Appli

33 3388 57.6 1161 3 US-09-193-043-2 Sequence 2, Appli

34 3388 57.6 1161 4 US-09-688-307A-2 Sequence 2, Appli

35 3388 57.6 1161 4 US-09-350-259-2 Sequence 2, Appli

36 3372.5 57.4 1161 1 US-08-485-618-99 Sequence 99, Appl

37 3372.5 57.4 1161 2 US-08-605-672-99 Sequence 99, Appl

38 3372.5 57.4 1161 2 US-08-482-293A-99 Sequence 99, Appl

39 3372.5 57.4 1161 2 US-08-943-363-99 Sequence 99, Appl

40 3372.5 57.4 1161 3 US-09-193-043-99 Sequence 99, Appl

41 3372.5 57.4 1161 4 US-09-688-307A-99 Sequence 99, Appl

42 3372.5 57.4 1161 4 US-09-350-259-99 Sequence 99, Appl

43 3213.5 54.7 1161 3 US-09-193-043-55 Sequence 55, Appl

44 3213.5 54.7 1161 4 US-09-688-307A-55 Sequence 55, Appl

45 3213.5 54.7 1161 4 US-09-350-259-55 Sequence 55, Appl

#### ALIGNMENTS

RESULT 1

US-08-173-497-3

; Sequence 3, Application US/08173497

; Patent No. 5437958

; GENERAL INFORMATION:

; APPLICANT: Gallatin, W. Michael

; APPLICANT: Van Der Vieren, Monica

; TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha

; TITLE OF INVENTION: Subunit

; NUMBER OF SEQUENCES: 29

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun

; STREET: 233 S. Wacker Drive, 6300 Sears Tower

; CITY: Chicago

; STATE: Illinois

; COUNTRY: USA

; ZIP: 60606-6402

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION NUMBER: US/08/173,497

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: No. 5437958and, Greta E.

; REGISTRATION NUMBER: 35,302

; REFERENCE/DOCKET NUMBER: 27866/31363

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 312-474-6300

; TELEFAX: 312-474-0448

; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1153 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-173-497-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 60

DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFGLFNSLRQPOK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFGLFNSLRQPOK 136  
QY 121 FPEARLGCQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMQYSEF 180  
DB 137 FPEARLGCQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMQYSEF 196  
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFTNGARKNAFKILFL 240  
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFTNGARKNAFKILV 256  
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNTVASKPRDHVFOAN 300  
DB 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNTVASKPRDHVFOQ 316  
QY 301 NFEALKTQNLQREKIFAEGTQGTSSSFHEMSQEGFSAATISNGPLSTVGSYDWAG 360  
DB 317 NFEALKTQNLQREKIFAEGTQGTSSSFHEMSQEGFSAATISNGPLSTVGSYDWAG 376  
QY 361 GVFLYTSKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 420  
DB 377 GVFLYTSKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYOHIGLVAMFR 436  
QY 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 480  
DB 437 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 496  
QY 481 PRGQARWOCDAVLXGEOQPGWFGAALTVDLVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGQARWOCDAVLXGEOQPGWFGAALTVDLVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFQOSLSGGQDLTMDGLVLDLVGAQHVLILRSQ 600  
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFQOSLSGGQDLTMDGLVLDLVGAQHVLILRSQ 616  
QY 601 PVLIRKAIKMEFPREVARNVFECNDQVVKGEAGEVVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLIRKAIKMEFPREVARNVFECNDQVVKGEAGEVVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSGRPHSAFVNETKSTRQTQVLGTCTETILKQLPNCIEDPVPSPVILRLNF 720  
DB 677 YDLALDSGRPHSAFVNETKSTRQTQVLGTCTETILKQLPNCIEDPVPSPVILRLNF 736  
QY 721 SLVGTPLSAFGNLRPVLAEDAQLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFGNLRPVLAEDAQLFTALPPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGSDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
DB 797 GPREFNVTVVRNDGSDSVRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 900  
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPTSLVFLVPV 960  
DB 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSIPTSLVFLVPV 976  
QY 961 RLNQTVIDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1020  
DB 977 RLNQTVIDRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDIP 1036  
QY 1021 FFGIQEFENATLKNLSFDWYIKTSHNHLIVSTABILFNDVSVFTLLPGOGAFVRSQTET 1080  
DB 1037 FFGIQEFENATLKNLSFDWYIKTSHNHLIVSTABILFNDVSVFTLLPGOGAFVRSQTET 1096  
QY 1081 KVEPFEPVNPPLPLIVGSSVGLLALLITAALYKLGFFKQYKDMMSGEGPPGAEQ 1137  
DB 1097 KVEPFEPVNPPLPLIVGSSVGLLALLITAALYKLGFFKQYKDMMSGEGPPGAEQ 1153

RESULT 2  
US-08-286-889-3  
; Sequence 3, Application US/08286889  
; Patent No. 5470953  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Mich  
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: P38,659  
; REFERENCE/DOCKET NUMBER: 27866/32168  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-286-889-3  
Query Match 99.3%; Score 5839; DB 1; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFGLFNSLRQPOK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVKGCLFGLFNSLRQPOK 136  
QY 121 FPEARLGCQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMQYSEF 180  
DB 137 FPEARLGCQSDSDIAFLIDGSGSIIPHDPRMKELVSTIMEQLKSKTFLSLMQYSEF 196  
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFTNGARKNAFKILFL 240  
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGKRVVRELFTNGARKNAFKILV 256  
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNTVASKPRDHVFOAN 300  
DB 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNTVASKPRDHVFOV 316  
QY 301 NFEALKTQNLQREKIFAEGTQGTSSSFHEMSQEGFSAATISNGPLSTVGSYDWAG 360  
DB 317 NFEALKTQNLQREKIFAEGTQGTSSSFHEMSQEGFSAATISNGPLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVLAGPRYOHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVLAGPRYOHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480  
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496  
QY 481 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
DB 557 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 616  
QY 601 PVLRVKAIMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLRVKAIMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSRPHRAVFNKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720  
DB 677 YDLALDSRPHRAVFNKSTRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPPEKCGNDNICODDLSTIFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFGLNRPVLAEDAQRFLTALFPPEKCGNDNICODDLSTIFSFMSLDCLVVG 796  
QY 781 GPRFNVTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWLACESASSTEV 840  
DB 797 GPRFNVTVTVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
DB 857 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVVMVTSHGVSPTKYNFTASENTSRVMOHQVSNLQSRSLPISLVELVVP 960  
DB 917 QLELPVKYAVVMVTSHGVSPTKYNFTASENTSRVMOHQVSNLQSRSLPISLVELVVP 976  
QY 961 RLNTQVTDWRPQVTFSENLSSTCHTKERLPSSHDFLAELKAPVNVNCSIAVCQRIQCDIP 1020  
DB 977 RLNTQVTDWRPQVTFSENLSSTCHTKERLPSSHDFLAELKAPVNVNCSIAVCQRIQCDIP 1036  
QY 1021 FFGQOEFNATLKNLSFDWIKTSHNELLIVSTAEILFNDVSVTLIPGQGAFFVRSQTE 1080  
DB 1037 FFGQOEFNATLKNLSFDWIKTSHNELLIVSTAEILFNDVSVTLIPGQGAFFVRSQTE 1096  
QY 1081 KVEPFEVNPPLPLTVGSSVGLLLALITLALYKLGFFKROYKDMSEGGPPGAEPQ 1137  
DB 1097 KVEPFEVNPPLPLTVGSSVGLLLALITLALYKLGFFKROYKDMSEGGPPGAEPQ 1153

## RESULT 3

US-08-485-618-3  
; Sequence 3, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485.618  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32797  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-485-618-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLGSRVVGAPQEIIVANORGLSYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFOENARFGQSVVQLGSRVVGAPQEIIVANORGLSYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTCSENTYVKGCLFLFGSNLRQPOK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTCSENTYVKGCLFLFGSNLRQPOK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQLKSKTLPFLMYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDFRMKELVSTIMEQLKSKTLPFLMYSEEF 196  
QY 181 RIHFTFKFONNPNRSLIKPIITOLLGRTHATGLRKVVRELFNITNGARKNAFKILL 240  
DB 197 RIHFTFKFONNPNRSLIKPIITOLLGRTHATGLRKVVRELFNITNGARKNAFKILL 256  
QY 241 TDGEKFGDPLGYEDVIPELDREGVIRVVLGFGDAFRSEKROELNTVASKPPRDHVFQAN 300  
DB 257 TDGEKFGDPLGYEDVIPELDREGVIRVVLGFGDAFRSEKROELNTVASKPPRDHVFQAN 316  
QY 301 NFPAKTVQNLREKIPAIETGTGSSSSFEHMSQGFSAATITNSGPELLSTVGSYDWAG 360  
DB 317 NFPAKTVQNLREKIPAIETGTGSSSSFEHMSQGFSAATITNSGPELLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVLAGPRYOHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVLAGPRYOHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480  
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496  
QY 481 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600



557 HGTSGSGISPHSQRISAGSKLSPLRQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 616  
601 PVLRVKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
617 PVLRVKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
661 YDLALDSGRPSRAVFNETHKSTRQTVGLGTQTCETLKLQLPNCIEDPSPVILRLNF 720  
677 YDLALDSGRPSRAVFNETHKSTRQTVGLGTQTCETLKLQLPNCIEDPSPVILRLNF 736  
721 SLVGTPLSAFAGNLRPVLAEADAQRLFTALFPFPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
737 SLVGTPLSAFAGNLRPVLAEADAQRLFTALFPFPEKNCNDNICODDLSITFSFMSLDCLVVG 796  
781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916  
901 QLELPVKYAVYVMVTSHGVSYKYLNFPTASENTSRVWQHGYQVSNLQSRSLPISLVFLVPV 960  
917 QLELPVKYAVYVMVTSHGVSYKYLNFPTASENTSRVWQHGYQVSNLQSRSLPISLVFLVPV 976  
961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPDSHDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020  
977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPDSHDFLAELRKAPVNVCSIAVCQRIQCDDIP 1036  
1021 PFGIOEEFNATLKGNSLDPWYIKTSNHLIYVSTAEILPNDSEVFTLLPQGGAPVRSQDET 1080  
1037 PFGIOEEFNATLKGNSLDPWYIKTSNHLIYVSTAEILPNDSEVFTLLPQGGAPVRSQDET 1096  
1081 KYPEPEVPNPLPLIVGSSVGGILLALITAAALYKLGFFKRRQYKDMWSEGGPPGAEPPQ 1137  
1097 KYPEPEVPNPLPLIVGSSVGGILLALITAAALYKLGFFKRRQYKDMWSEGGPPGAEPPQ 1153

RESULT 4  
US-08-362-652-3  
; Sequence 3, Application US/08362652  
; Patent No. 5766850  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,652  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-362-652-3

Query Match 99.3%; Score 5839; DB 1; Length 1153;  
Beat Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFQGSVVQIQGSRVVVVGAPQEIIVAAANQSGLYQCDYSTGSCBPI 60  
DB 17 FNLDTENAMTFOENARGFQGSVVQIQGSRVVVVGAPQEIIVAAANQSGLYQCDYSTGSCBPI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLFGLNLRQPOK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLFGLNLRQPOK 136  
QY 121 FPEALRGCPQEDSDTAFILIDSGSIIPHDFRMKELVSTIMEOLKSKTLFSLIMOYSBEP 180  
DB 137 FPEALRGCPQEDSDTAFILIDSGSIIPHDFRMKELVSTIMEOLKSKTLFSLIMOYSBEP 196  
QY 181 RHFTFKFQFONNPNRSLIKPITQLLGRTHATGLRKVVRELFTNNGARKNAKILFL 240  
DB 197 RHFTFKFQFONNPNRSLIKPITQLLGRTHATGLRKVVRELFTNNGARKNAKILFL 256  
QY 241 TDGEKFGDPLGYEDVIPLEDRGVIRYVLFQGDAPRSEKSRQELMTVASKPRPDHVFQAN 300  
DB 257 TDGEKFGDPLGYEDVIPLEDRGVIRYVLFQGDAPRSEKSRQELMTVASKPRPDHVFQAN 316  
QY 301 NFEALKTQVQNLREKIFATEGTQTSSESEFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360  
DB 317 NFEALKTQVQNLREKIFATEGTQTSSESEFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILRNVRQSLVLGAPRYQHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAAIILRNVRQSLVLGAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEQTREGQVSVCP 480  
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEQTREGQVSVCP 496  
QY 481 PRGQARWQCDVLYGEGQGPWGRGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGQARWQCDVLYGEGQGPWGRGAALTVLGDVNGDKLTVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPHSQRISAGSKLSPLRQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 600  
DB 557 HGTSGSGISPHSQRISAGSKLSPLRQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 616  
QY 601 PVLRVKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLRVKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSGRPSRAVFNETHKSTRQTVGLGTQTCETLKLQLPNCIEDPSPVILRLNF 720  
DB 677 YDLALDSGRPSRAVFNETHKSTRQTVGLGTQTCETLKLQLPNCIEDPSPVILRLNF 736  
QY 721 SLVGTPLSAFAGNLRPVLAEADAQRLFTALFPFPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFAGNLRPVLAEADAQRLFTALFPFPEKNCNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840

Db 797 GPRENVTVTVRNDGEDSVYRTQVTFPPPLDLISYKRVSTLQNRQSRKSWELACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIFFPENSEVFNITFDVDSKASLGKLLKLLKANVTSENNMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIFFPENSEVFNITFDVDSKASLGKLLKLLKANVTSENNMPRTNKTEF 916  
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLRPLSLVFLVPV 960  
Db 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLRPLSLVFLVPV 976  
Qy 961 RLNQTVIWDNRQVTPSENLSSTCHTKERLPKSHDSFLAELRKAPVNVNCIAVCQRIQCDDIP 1020  
Db 977 RLNQTVIWDNRQVTPSENLSSTCHTKERLPKSHDSFLAELRKAPVNVNCIAVCQRIQCDDIP 1036  
Qy 1021 FFGIQEENATLKGNSLDWIKTSHNHLIVSTAEILLFNDVSFVTLFPQOGAFVRSOTET 1080  
Db 1037 FFGIQEENATLKGNSLDWIKTSHNHLIVSTAEILLFNDVSFVTLFPQOGAFVRSOTET 1096  
Qy 1081 KVEPPEVENPLPLIVGSSVGGILLALITALYKLGFFKROVKDMWSEGGPPGAEPPQ 1137  
Db 1097 KVEPPEVENPLPLIVGSSVGGILLALITALYKLGFFKROVKDMWSEGGPPGAEPPQ 1153

## RESULT 5

US-08-605-672-3  
; Sequence 3, Application US/08605672  
; Patent No. 5817515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; City: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/605.672  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-605-672-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 ENLDTENAMTQENARGFGQSVVOLQGRVVVGAPOEIVAAORGLSYQCXYSTSCSCEPI 60  
Db 17 ENLDTENAMTQENARGFGQSVVOLQGRVVVGAPOEIVAAORGLSYQCXYSTSCSCEPI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLCFGLSGNLQOQPOK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLCFGLSGNLQOQPOK 136  
Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDPRMKELVSTIMEQLKKSKTFLSMOYSEEF 180  
Db 137 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDPRMKELVSTIMEQLKKSKTFLSMOYSEEF 196  
Qy 181 RIHFTFEFQNNPNRSLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILVVI 240  
Db 197 RIHFTFEFQNNPNRSLIKPITQLLGRTHATGLRKVRELFNITNGARKNAFKILVVI 256  
Qy 241 TDGEKFGDPLGYEDVIPEDLREGVIRYVVGFDAPRSEKSRQELNTVASKPRDRHVFQAN 300  
Db 257 TDGEKFGDPLGYEDVIPEDLREGVIRYVVGFDAPRSEKSRQELNTVASKPRDRHVFQAN 316  
Qy 301 NFEALKTIVONOLREKIFAIEGTQTGSSSSPHEMSQEGFSAITNSGPLLSTVGSYDWAG 360  
Db 317 NFEALKTIVONOLREKIFAIEGTQTGSSSSPHEMSQEGFSAITNSGPLLSTVGSYDWAG 376  
Qy 361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGAAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVSDMDNDAYLGAAAAIILNRVQSLVGLAPRYOHIGLVAMFR 436  
Qy 421 QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSITDLVLIGAPHYEQTRGGQVSVCP 480  
Db 437 QNTGMWESNANVKCTQIGAYFGASLCSVDVDSNGSITDLVLIGAPHYEQTRGGQVSVCP 496  
Qy 481 PRGQRAWQCDVLYGEGQGPWGFAGALTVLGVNGDKLTDVAIGAPGEDNNGAVLYF 540  
Db 497 PRGQRAWQCDVLYGEGQGPWGFAGALTVLGVNGDKLTDVAIGAPGEDNNGAVLYF 556  
Qy 541 HGTSGSGISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTGAGHVLRLSRQ 600  
Db 557 HGTSGSGISPSHSORIASGLSPRLQYFGOSLGGQDLTMDGLVDLTGAGHVLRLSRQ 616  
Qy 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 617 PVLRVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRVAVFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720  
Db 677 YDLALDSGRPHSRVAVFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736  
Qy 721 SLVGTPLSAFNLRLPVLAEADAQRLFTALFPFEKNCNDNICODDLSITFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLRLPVLAEADAQRLFTALFPFEKNCNDNICODDLSITFMSLDCLVVG 796  
Qy 781 GPREFNVTVTVRNDGEDSVYRTQVTFPPPLDLISYKRVSTLQNRQSRKSWELACESASSTEV 840  
Db 797 GPREFNVTVTVRNDGEDSVYRTQVTFPPPLDLISYKRVSTLQNRQSRKSWELACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIFFPENSEVFNITFDVDSKASLGKLLKLLKANVTSENNMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIFFPENSEVFNITFDVDSKASLGKLLKLLKANVTSENNMPRTNKTEF 916  
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLRPLSLVFLVPV 960  
Db 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLRPLSLVFLVPV 976  
Qy 961 RLNQTVIWDNRQVTPSENLSSTCHTKERLPKSHDSFLAELRKAPVNVNCIAVCQRIQCDDIP 1020  
Db 977 RLNQTVIWDNRQVTPSENLSSTCHTKERLPKSHDSFLAELRKAPVNVNCIAVCQRIQCDDIP 1036

QY 1021 PFGIOEENATLKGNSLSDWYKTSNHLIIIVSTAEIILFNDVSFTLLPGQGAFFVSOTET 1080  
DB 1037 PFGIOEENATLKGNSLSDWYKTSNHLIIIVSTAEIILFNDVSFTLLPGQGAFFVSOTET 1096  
QY 1081 KVEPEVENPPLIIVGSSVGGILLALITAAALYKLGFFKROKYMMSGGPPGABPQ 1137  
DB 1097 KVEPEVENPPLIIVGSSVGGILLALITAAALYKLGFFKROKYMMSGGPPGABPQ 1153

RESULT 6  
US-08-482-293A-3  
; Sequence 3, Application US/08482293A  
; Patent No. 5831029  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,293A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-482-293A-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPT 60  
DB 17 ENLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPT 76  
QY 61 RLQVPEAVNMISGLSLAATSPQLIACGPTVHOTCSENTYVYKGLCFLFGSNLRQOPQK 120  
DB 77 RLQVPEAVNMISGLSLAATSPQLIACGPTVHOTCSENTYVYKGLCFLFGSNLRQOPQK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQLKSKTFLSLMQSYSEF 180

DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQLKSKTFLSLMQSYSEF 196  
QY 191 RIHFTFKFQNNPNSRLIKITQLLGRTHATGLRKVVRELFINITGNARKNAFLIFLL 240  
DB 197 RIHFTFKFQNNPNSRLIKITQLLGRTHATGLRKVVRELFINITGNARKNAFLIVVI 256  
QY 241 TDGEKFGDPLGVEDVPELDRGVIRYVLFGDAPRSKRSQELMTVASKPRDPRHVFQAN 300  
DB 257 TDGEKFGDPLGVEDVPEADREGVIRYVIGVDDAPRSKRSQELMTIASKPRDPRHVFQV 316  
QY 301 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGPLLSITVGYDWAG 360  
DB 317 NFEALKTQNLQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSNGPLLSITVGYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVLGAPRYQHIGLVAMER 420  
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVLGAPRYQHIGLVAMER 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQSVVCP 480  
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYEQTRGGQSVVCP 496  
QY 481 PRGQARWQCDVAVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGQARWQCDVAVLYGEOQPMGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVLDITVCAQGHVLLLRQ 600  
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVLDITVCAQGHVLLLRQ 616  
QY 601 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 660  
DB 617 PVLRYKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGTQTCETLKLQPCIEDPVPFIVLRNPF 720  
DB 677 YDLALDSGRPHSRVAFNETKNSRRQTQVLGTQTCETLKLQPCIEDPVPFIVLRNPF 736  
QY 721 SLVGTPLSAFNLRLPVLAEADQRLFTALPPFEKNCNDNICODDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFNLRLPVLAEADQRLFTALPPFEKNCNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLGYRKVSTLQONORSQSWELACESASSTEV 840  
DB 797 GPREFNVTVVRNDGEDSVRTQVTFPPFLDLGYRKVSTLQONORSQSWELACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 900  
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYLAFTASENTSRVMQHOYQVSNLQORSILPISLVLPV 960  
DB 917 QLELPVKYAVYVMTSHGVSTKYLAFTASENTSRVMQHOYQVSNLQORSILPISLVLPV 976  
QY 961 RLNQTVIMDRPQVTFSENLSSSTCHTKERLPSHSDFLAELKAPVNVNCSTAVCQRIQCDIP 1020  
DB 977 RLNQTVIMDRPQVTFSENLSSSTCHTKERLPSHSDFLAELKAPVNVNCSTAVCQRIQCDIP 1036  
QY 1021 PFGIOEENATLKGNSLSDWYKTSNHLIIIVSTAEIILFNDVSFTLLPGQGAFFVSOTET 1080  
DB 1037 PFGIOEENATLKGNSLSDWYKTSNHLIIIVSTAEIILFNDVSFTLLPGQGAFFVSOTET 1096  
QY 1081 KVEPEVENPPLIIVGSSVGGILLALITAAALYKLGFFKROKYMMSGGPPGABPQ 1137  
DB 1097 KVEPEVENPPLIIVGSSVGGILLALITAAALYKLGFFKROKYMMSGGPPGABPQ 1153

RESULT 7  
US-08-943-363-3  
; Sequence 3, Application US/08943363  
; Patent No. 5837478  
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943.363  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-943-363-3

Query Match 99.3%; Score 5839; DB 2; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGQSVVLOGSRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARFGQSVVLOGSRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVAVNMSGLSLAATTSPPQLAACGPTVHQTCSNTYVKGCLFLGSLNLRQOPQK 120  
Db 77 RLQVPVAVNMSGLSLAATTSPPQLAACGPTVHQTCSNTYVKGCLFLGSLNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLDGSGSIIPDPRRMKELVSTIMEQLKSKTLFSLMQXSEEF 180  
Db 137 FPEALRGCPQEDSDIAFLDGSGSIIPDPRRMKELVSTIMEQLKSKTLFSLMQXSEEF 196

QY 181 RIHFTKFEQNNPRSLIKPIQLLGRTHATGLRWKVELNITNGARKNAFKILFLL 240  
Db 197 RIHFTKFEQNNPRSLIKPIQLLGRTHATGLRWKVELNITNGARKNAFKILFLL 256

QY 241 TDGEKGDPLGYEDVPEADREGVIRVVGDAFSEKSKQELNTVASKPPRDHVFQAN 300  
Db 257 TDGEKGDPLGYEDVPEADREGVIRVVGDAFSEKSKQELNTVASKPPRDHVFQAN 316

QY 301 NFALKTVQNLREKIFAIBGTQTGSSSSFEHMSQEGFSAATSNGLPSTVGSYDWAG 360  
Db 317 NFALKTVQNLREKIFAIBGTQTGSSSSFEHMSQEGFSAATSNGLPSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTGTGQGVSVCLP 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTGTGQGVSVCLP 496

QY 481 PRGQARWQCDAVLYGEOGQPGWRFGAALTVDLGDVNGDKLTDVAIGAPEERNGAVILF 540  
Db 497 PRGQARWQCDAVLYGEOGQPGWRFGAALTVDLGDVNGDKLTDVAIGAPEERNGAVILF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAGHQHVLLLSRQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAGHQHVLLLSRQ 616

QY 601 PVLRVKATMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 617 PVLRVKATMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720  
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736

QY 721 SLVGTPLSAFONLRPVLAEADAQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFONLRPVLAEADAQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796

QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYKRVSTLQNRQSRWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYKRVSTLQNRQSRWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916

QY 901 QLELPVKYAVVMVTSHGVTSKYLNFNTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 960  
Db 917 QLELPVKYAVVMVTSHGVTSKYLNFNTASENTSRVMOHQYQVSNLQORSIPISLVFLVPV 976

QY 961 RLNQTVIWDPRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
Db 977 RLNQTVIWDPRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036

QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1080  
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSOTET 1096

QY 1081 KVEPFEVNPPLIIVGSSVGLLLALLITAALYKLGFFKQYKDMMSSEGGPPGAEPO 1137  
Db 1097 KVEPFEVNPPLIIVGSSVGLLLALLITAALYKLGFFKQYKDMMSSEGGPPGAEPO 1153

RESULT 8  
US-09-193-043-3  
; Sequence 3, Application US/09193043  
; Patent No. 6251395  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6251395el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/193,043  
; CURRENT FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: Patent in Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-193-043-3									
Query Match									
Best Local Similarity 99.3%; Score 5839; DB 3; Length 1153;									
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;									
Qy	1	ENLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCBPI	60						
Db	17	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCBPI	76						
Qy	61	RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRQPOK	120						
Db	77	RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRQPOK	136						
Qy	121	FPFALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLSLMOYSSEF	180						
Db	137	FPFALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLSLMOYSSEF	196						
Qy	181	RHFTEKFEQNNPNRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAFKILFL	240						
Db	197	RIHFTKFEQNNPNRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAFKILV	256						
Qy	241	TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDRHVQAN	300						
Db	257	TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDRHVQAN	316						
Qy	301	NFEALKTIVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG	360						
Db	317	NFEALKTIVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG	376						
Qy	361	GVFLYTSKESKFINNTRVDSMDNDAYLGAAAILNRVQSLVILGAPYQHIGLVAMPR	420						
Db	377	GVFLYTSKESKFINNTRVDSMDNDAYLGAAAILNRVQSLVILGAPYQHIGLVAMPR	436						
Qy	421	QNTGMWESNANVKGTOIGAFYASLCSVDVDSNGSTDVLVILGAPHYEQTRGGQSVCP	480						
Db	437	QNTGMWESNANVKGTOIGAFYASLCSVDVDSNGSTDVLVILGAPHYEQTRGGQSVCP	496						
Qy	481	PRQGRARWQCDVLYGEGQGPWGRFGAALTIVLDVNGDKLTDAIAGPGEEDNRGAVILF	540						
Db	497	PRQGRARWQCDVLYGEGQGPWGRFGAALTIVLDVNGDKLTDAIAGPGEEDNRGAVILF	556						
Qy	541	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTGAAQGHVLLRSQ	600						
Db	557	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTGAAQGHVLLRSQ	616						
Qy	601	PVLVRKAIMFENPREVARNVFECDQVWKGKEAGEVRVCLHVKQSTRDLRREGQIOSVVT	660						
Db	617	PVLVRKAIMFENPREVARNVFECDQVWKGKEAGEVRVCLHVKQSTRDLRREGQIOSVVT	676						
Qy	661	YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETIKLQLPNCIEDPVSIVILRLNF	720						
Db	677	YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETIKLQLPNCIEDPVSIVILRLNF	736						
Qy	721	SLVGTPLSAFNLRPVLAEDAQLFTALPFEKNCGNDNICDDLSITFSFMSLDCLVVG	780						
Db	737	SLVGTPLSAFNLRPVLAEDAQLFTALPFEKNCGNDNICDDLSITFSFMSLDCLVVG	796						
Qy	781	GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLACSSASSTEV	840						
Db	797	GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNSQRSWRLACSSASSTEV	856						
Qy	841	SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF	900						
Db	857	SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF	916						
Qy	901	QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLPISLVFLVPV	960						

Db	917	QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLPISLVFLVPV	976						
Qy	961	RLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP	1020						
Db	977	RLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP	1036						
Qy	1021	FFGIQEEFNATLKGNLGDFDWYIKTSHNHLIVSTAELFNDVSFTLLPGQGAFAVRSQDET	1080						
Db	1037	FFGIQEEFNATLKGNLGDFDWYIKTSHNHLIVSTAELFNDVSFTLLPGQGAFAVRSQDET	1096						
Qy	1081	KVBEPFVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKQKMMSEGGPPGABEQ	1137						
Db	1097	KVBEPFVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKQKMMSEGGPPGABEQ	1153						
RESULT 9									
US-09-688-307A-3									
; Sequence 3, Application US/09688307A									
; Patent No. 6432404									
; GENERAL INFORMATION:									
; APPLICANT: Gallatin, Michael W.									
; APPLICANT: Van der Vieren, Monica									
; TITLE OF INVENTION: No. 6432404el Human Beta-2									
; FILE REFERENCE: 278667/36646									
; CURRENT APPLICATION NUMBER: US/09/688,307A									
; CURRENT FILING DATE: 2000-10-13									
; PRIOR APPLICATION NUMBER: 09/193,043									
; PRIOR FILING DATE: 1998-11-16									
; PRIOR APPLICATION NUMBER: 08/605,672									
; PRIOR FILING DATE: 1996-02-22									
; PRIOR APPLICATION NUMBER: 08/173,497									
; PRIOR FILING DATE: 1993-12-23									
; PRIOR APPLICATION NUMBER: 08/286,889									
; PRIOR FILING DATE: 1994-08-05									
; PRIOR APPLICATION NUMBER: 08/362,652									
; PRIOR FILING DATE: 1994-12-21									
; PRIOR APPLICATION NUMBER: 08/943,363									
; PRIOR FILING DATE: 1997-10-03									
; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: Patent in Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-688-307A-3									
Query Match									
Best Local Similarity 99.3%; Score 5839; DB 4; Length 1153;									
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;									
Qy	1	ENLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCBPI	60						
Db	17	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCBPI	76						
Qy	61	RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRQPOK	120						
Db	77	RLQVPEAVNMSLGLSLAATSPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRQPOK	136						
Qy	121	FPFALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLSLMOYSSEF	180						
Db	137	FPFALRGCPQEDSDIAFLIDGSGSIIPHDFRMKELVSTIMEOLKSKTFLSLMOYSSEF	196						
Qy	181	RHFTEKFEQNNPNRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAFKILFL	240						
Db	197	RIHFTKFEQNNPNRSLIKPITQLGRTHATGLRKVVRELFNITNGARKNAFKILV	256						
Qy	241	TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDRHVQAN	300						
Db	257	TDGEKFGDPLGYEDVIPEDREGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDRHVQAN	316						
Qy	301	NFEALKTIVQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG	360						

Db 317 NFEALKTIONQUREKIFAIEGTQGTGSSSSFEHMSQEGFSAATTSNGPLLSITVGSYDWAG 376  
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAATILRNREVQSILVILGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAATILRNREVQSILVILGAPRYQHIGLVAMFR 436  
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVILGAPHYIEQTRGGQVSVCP 480  
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVILGAPHYIEQTRGGQVSVCP 496  
Qy 481 PRGORARWQCDVILYGEQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGORARWQCDVILYGEQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
Qy 601 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660  
Db 617 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRVAVNETKNSSTRQTQVILGLTQTCETLKLQLPNCIEDPVPVILRLNF 720  
Db 677 YDLALDSGRPHSRVAVNETKNSSTRQTQVILGLTQTCETLKLQLPNCIEDPVPVILRLNF 736  
Qy 721 SLVGTPLSAFNLRPVLAEDAQRFTALPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSNNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSNNMPTNKTEF 916  
Qy 901 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHOYVSNLQORSLSILVFLVPV 960  
Db 917 QLELPVKYAVVMVTSHGVS TKYLNFTASENTSRVMQHOYVSNLQORSLSILVFLVPV 976  
Qy 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIACVQRIQCDIP 1020  
Db 977 RLNQTIVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCVSIACVQRIQCDIP 1036  
Qy 1021 FFGIQEFENATLKNLSPDWIKTSHNHLILVSTABILFNDVSFTLLPGQGFVRSQTET 1080  
Db 1037 FFGIQEFENATLKNLSPDWIKTSHNHLILVSTABILFNDVSFTLLPGQGFVRSQTET 1096  
Qy 1081 KVEPFEPVNPPLIVGSSVGLLLALITAALYKLGFKKRYKDMSEGGPGBEPQ 1137  
Db 1097 KVEPFEPVNPPLIVGSSVGLLLALITAALYKLGFKKRYKDMSEGGPGBEPQ 1153

## RESULT 10

US-09-350-259-3  
; Sequence 3, Application US/09350259  
; Patent No. 6620915  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6620915el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350, 259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-3

Query Match 99.3%; Score 5839; DB 4; Length 1153;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 8; Mismatches 5; Indels 0; Gaps 0;  
Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQORSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANQORSLYQCDYSTGSCPEI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTQCSNTYVVKGLCFGLGSNLRQPOK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHTQCSNTYVVKGLCFGLGSNLRQPOK 136  
Qy 121 FPEALRGCPQSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTILSLMOYSEEP 180  
Db 137 FPEALRGCPQSDIAFLIDGSGSIIPHDPRMKELVSTIMEOLKSKTILSLMOYSEEP 196  
Qy 181 RIHFTFEFQNNPNRSLIKPITOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240  
Db 197 RIHFTFEFQNNPNRSLIKPITOLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 256  
Qy 241 TDGKFKGDPGLGYEDVIBELDREGVIRVYVGLFGDAPRSEKSOELNTVASKPRDRHVQAN 300  
Db 257 TDGKFKGDPGLGYEDVIBELDREGVIRVYVGLFGDAPRSEKSOELNTVASKPRDRHVQAN 316  
Qy 301 NFEALKTVQNLREKIIFAIEGTQGTGSSSSFEHMSQEGFSAATTSNGPLLSITVGSYDWAG 360  
Db 317 NFEALKTVQNLREKIIFAIEGTQGTGSSSSFEHMSQEGFSAATTSNGPLLSITVGSYDWAG 376  
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAATILRNREVQSILVILGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAATILRNREVQSILVILGAPRYQHIGLVAMFR 436  
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVILGAPHYIEQTRGGQVSVCP 480  
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLVILGAPHYIEQTRGGQVSVCP 496  
Qy 481 PRGORARWQCDVILYGEQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGORARWQCDVILYGEQGPWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
Qy 601 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660  
Db 617 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRVAVNETKNSSTRQTQVILGLTQTCETLKLQLPNCIEDPVPVILRLNF 720  
Db 677 YDLALDSGRPHSRVAVNETKNSSTRQTQVILGLTQTCETLKLQLPNCIEDPVPVILRLNF 736  
Qy 721 SLVGTPLSAFNLRPVLAEDAQRFTALPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALPPEKNCNGNDNICODDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSNNMPTNKTEF 900

Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 916  
Qy 901 QLELPVKYAVYVMVTSYHGVSTKYLNFNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960  
Db 917 QLELPVKYAVYVMVTSYHGVSTKYLNFNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 976  
Qy 961 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPKSHSDFLAELRKAPVNCSTAVCORIQCDIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPKSHSDFLAELRKAPVNCSTAVCORIQCDIP 1036  
Qy 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQCGAFVRSQDET 1080  
Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQCGAFVRSQDET 1096  
Qy 1081 KVEPEVPNPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSRGPPGABEQ 1137  
Db 1097 KVEPEVPNPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSRGPPGABEQ 1153

RESULT 11

US-08-476-062A-43  
; Sequence 43, Application US/08476062A  
; Patent No. 5877275

GENERAL INFORMATION:

; APPLICANT: Arnaout, M. Amin  
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY  
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

; NUMBER OF SEQUENCES: 53

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: US

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/476,062A

; FILING DATE: 07-JUN-1995

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: 08/216,081

; FILING DATE: 21-MAR-1994

; APPLICATION NUMBER: 07/637,830

; FILING DATE: 04-JAN-1991

; APPLICATION NUMBER: 07/539,842

; FILING DATE: 18-JUN-1990

; APPLICATION NUMBER: 07/212,573

; FILING DATE: 28-JUN-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Freeman, John W.

; REGISTRATION NUMBER: 29,066

; REFERENCE/DOCKET NUMBER: 00786/068003

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1152 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

; US-08-476-062A-43

Query Match 98.8%; Score 5808.5; DB 2; Length 1152;  
Best Local Similarity 98.6%; Pred. No. 0;  
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTPOENARGFGQSVVQLOGSRVVVGAPQEIIVAAANQRGSLVQCYSTGSCBPI 60  
Db 17 FNLDTENAMTPOENARGFGQSVVQLOGSRVVVGAPQEIIVAAANQRGSLVQCYSTGSCBPI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGTVHQTCSNTYVXGLCLFLGSLNRQQPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGTVHQTCSNTYVXGLCLFLGSLNRQQPQK 136  
Qy 121 FPEALRGCPQSDIDIAELIDGSGSIIIPDFFRMKELVSTIMBOLKSKSTLFSLMQSYBEF 180  
Db 137 FPEALRGCPQSDIDIAELIDGSGSIIIPDFFRMKEFVSTVMEQLKSKSTLFSLMQSYBEF 196  
Qy 181 RIHFTFKFQNNPNRSLIK3ITQLLGRTHATGRLKRVVRELFNTNGARKNAFKILFLL 240  
Db 197 RIHFTFKFQNNPNRSLIK3ITQLLGRTHATGRLKRVVRELFNTNGARKNAFKILVVI 256  
Qy 241 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHVFOAN 300  
Db 257 TDGEKFGDPLGYEDVIPLEADREGVIRYVIGVGDAFRSEKSRQELNTVASKPRDHVFOAN 316  
Qy 301 NPEALKTVQNLREKIFAIEGTQTGSSSSSFHEMSQEGFSAAITSNGLPSTLVGSYDNAG 360  
Db 317 NPEALKTVQNLREKIFAIEGTQTGSSSSSFHEMSQEGFSAAITSNGLPSTLVGSYDNAG 376  
Qy 361 GYFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMPR 420  
Db 377 GYFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMPR 436  
Qy 421 QNTGWESNANVKGTOIGAYFGASICSDVDVDSNGSTDLVLGAPHYETRGQGVSCPL 480  
Db 437 QNTGWESNANVKGTOIGAYFGASICSDVDVDSNGSTDLVLGAPHYETRGQGVSCPL 496  
Qy 481 PRGQARWOCDAVLYGEOQPMGRFGAALTVDLVNGDKLTDVAIGAPGEENRGAVYLF 540  
Db 497 PRG-RARWOCDAVLYGEOQPMGRFGAALTVDLVNGDKLTDVAIGAPGEENRGAVYLF 555  
Qy 541 HGTSGSIGSPSHSQRISAGKLSPLRQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
Db 556 HGTSGSIGSPSHSQRISAGKLSPLRQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 615  
Qy 601 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 660  
Db 616 PVLRVKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIOSVVT 675  
Qy 661 YDLALDSGRPHGRAVENETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSPIVLRNF 720  
Db 676 YDLALDSGRPHGRAVENETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSPIVLRNF 735  
Qy 721 SLVGTPLSAFNLRLPVLAEADAQRLFTALPPFEKNCGNDNICQDDLSITPFSFMSLDCLVVG 780  
Db 736 SLVGTPLSAFNLRLPVLAEADAQRLFTALPPFEKNCGNDNICQDDLSITPFSFMSLDCLVVG 795  
Qy 781 GPREFNVTYVRNDGEDSVRTQVTFPPFLDLSYRKVSTIQNORSQSRWLACESASSTEV 840  
Db 796 GPRESNVTYVRNDGEDSVRTQVTFPPFLDLSYRKVSTIQNORSQSRWLACESASSTEV 855  
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900  
Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 915  
Qy 901 QLELPVKYAVYVMVTSYHGVSTKYLNFNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 960  
Db 916 QLELPVKYAVYVMVTSYHGVSTKYLNFNTASNTSRVMOHQYQVSNLQORSLSPLSLVFLVPV 975  
Qy 961 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPKSHSDFLAELRKAPVNCSTAVCORIQCDIP 1020  
Db 976 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPKSHSDFLAELRKAPVNCSTAVCORIQCDIP 1035  
Qy 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQCGAFVRSQDET 1080  
Db 1036 PFGIOEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQCGAFVRSQDET 1095  
Qy 1081 KVEPEVPNPLIIVGSSVGGILLALITAAALYKLGFFKQYKDMMSRGPPGABEQ 1137



```

Db 1096 KVEPPEVENPLPLIVGSSVGGLLALLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1152
|||||
RESULT 12
PCT-US96-01314-43
; Sequence 43, Application PC/TUS9601314
; GENERAL INFORMATION:
; APPLICANT: M. Amin Arnaut
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN
; TITLE OF INVENTION: ANTAGONISTS
; NUMBER OF SEQUENCES: 78
; CORRESPONDENCE ADDRESS:
; ADDRESS: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/01314
; FILING DATE: 30-JAN-96
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/380,167
; FILING DATE: 30-JAN-95
; ATTORNEY/AGENT INFORMATION:
; NAME: John W. Freeman
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 00786/267001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1152
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US96-01314-43

Query Match 98.8%; Score 5808.5; DB 5; Length 1152;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQIIVANQSGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQIIVANQSGSLYQCDYSTGSCPEI 76
Qy 61 RLQVPVAVNMSGLSLAATTSPPQLACGPTVHTQCSNTYVKGCLCFGLGSLNRQOPQK 120
Db 77 RLQVPVAVNMSGLSLAATTSPPQLACGPTVHTQCSNTYVKGCLCFGLGSLNRQOPQK 136
Qy 121 FPEALRCPOEDSDIAPLDGSGSIIHPDFRMKELVSTTMEOLKKGKTTIFSLMQYSEEF 180
Db 137 FPEALRCPOEDSDIAPLDGSGSIIHPDFRMKELVSTTMEOLKKGKTTIFSLMQYSEEF 196
Qy 181 RIHFTFEFQNNPNRSLKPIITQLGRTHATGLRVKVRVRELNITNGARKNAFKILFL 240
Db 197 RIHFTFEFQNNPNRSLKPIITQLGRTHATGLRVKVRVRELNITNGARKNAFKILVI 256
Qy 241 TDEKFGDPIGYEDVPELDEGVIRYVLGFGDAFRSEKSRQELNTVASKPPRDHVFQAN 300
Db 257 TDEKFGDPIGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQV 316
Qy 301 NFBAKLTQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAATTSNGPLLSTVGSYDWAG 360
Db 317 NFBAKLTQNLREKIFAIEGTQTGSSSSFEHEMSQEGFSAATTSNGPLLSTVGSYDWAG 376

```

```

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTROGQSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTROGQSVCP 496
Qy 481 PRGQARWQCDVILYGEQGPWGRFGAALTVDLGVNGDKLTDVAICAPGEEDNRGAIVL 540
Db 497 PRG-RARWQCDVILYGEQGPWGRFGAALTVDLGVNGDKLTDVAICAPGEEDNRGAIVL 555
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAQSHVLLRSQ 600
Db 556 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAQSHVLLRSQ 615
Qy 601 PVLRVKALMEFPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
Db 616 PVLRVKALMEFPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 675
Qy 661 YDLALDSGRPHSRAVFNETKNSTRRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 676 YDLALDSGRPHSRAVFNETKNSTRRQTOVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 735
Qy 721 SLVGTPLSAFNGLRPVLAEDAQRIFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 736 SLVGTPLSAFNGLRPVLAEDAQRIFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 795
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 796 GPRESNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 915
Qy 901 QLELPVKYAVYVMTSHGVSTKYINFTASNTSRVMOHQYQVSNLQSRUPISLVFLVPV 960
Db 916 QLELPVKYAVYVMTSHGVSTKYINFTASNTSRVMOHQYQVSNLQSRUPISLVFLVPV 975
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1020
Db 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCIAVCQRIQCDIP 1035
Qy 1021 FFGIQEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
Db 1036 FFGIQEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1095
Qy 1081 KVEPPEVNPPLPLIVGSSVGGLLALLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1137
Db 1096 KVEPPEVNPPLPLIVGSSVGGLLALLALITAAALYKLGFFKQYKDMWSEGGPPGAEPQ 1152

```

```

RESULT 13
5424399-2
; Patent No. 5424399
; APPLICANT: ARNAOUT, M. AMIN
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS
; NUMBER OF SEQUENCES: 12
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/78,871
; FILING DATE: 16-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 539,842
; FILING DATE: 18-JUN-1990
; APPLICATION NUMBER: 212,573
; FILING DATE: 28-JUN-1988
; SEQ ID NO: 2
; LENGTH: 1152
5424399-2

```

Query Match 98.8%; Score 5808.5; DB 6; Length 1152;

Best Local Similarity 98.6%; Pred. No. 0;  
Matches 1121; Conservative 8; Mismatches 7; Indels 1; Gaps 1;

QY 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPEIVAANQSGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPEIVAANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPPQLLAGCPVHTQCSNTYVVKGLCLFLGSLNLRQOQPK 120  
Db 77 RLQVPEAVNMSLGLSLAATTSPPQLLAGCPVHTQCSNTYVVKGLCLFLGSLNLRQOQPK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEOLKSKTFLSLMOYSEEF 180  
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEOLKSKTFLSLMOYSEEF 196

QY 181 RIHFTFKEFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFTNGARKNAFKILFLL 240  
Db 197 RIHFTFKEFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFTNGARKNAFKILFLL 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVGFDAFRSEKSRQELANTVASKPRDHVFOAN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVGFDAFRSEKSRQELANTVASKPRDHVFOAN 316

QY 301 NFEALKTQONLREKIFALEGTGTGSSSFEHMSQEGFSAATISNGPLLSVTGSDYDAG 360  
Db 317 NFEALKTQONLREKIFALEGTGTGSSSFEHMSQEGFSAATISNGPLLSVTGSDYDAG 376

QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVIGAPRYQHIGLVAMPR 420  
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAAIILNRVQSLVIGAPRYQHIGLVAMPR 436

QY 421 QNTGHWESNANVKGTOIGAYFGASLCSDVDNSGTDLVIGAPHYVETRQGVSVCP 480  
Db 437 QNTGHWESNANVKGTOIGAYFGASLCSDVDNSGTDLVIGAPHYVETRQGVSVCP 496

QY 481 PRGQARWOCDAVLYEQOQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWOCDAVLYEQOQPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

QY 541 HTGSSGSIIPSHSQRISAGSLPRLQYFGQSLGGQDLTMDGLDVLTVGAQGHVLLRSQ 600  
Db 556 HTGSSGSIIPSHSQRISAGSLPRLQYFGQSLGGQDLTMDGLDVLTVGAQGHVLLRSQ 615

QY 601 PVLKVAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGIOQSVVT 660  
Db 616 PVLKVAIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGIOQSVVT 675

QY 661 YDLALDSGRPHSRAVNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720  
Db 676 YDLALDSGRPHSRAVNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 735

QY 721 SLVGTPLSAFNLRPVLAEDAQRLFTALPFPFNKCGNDNICQDDLITFTSFMSLDCLVVG 780  
Db 736 SLVGTPLSAFNLRPVLAEDAQRLFTALPFPFNKCGNDNICQDDLITFTSFMSLDCLVVG 795

QY 781 GPREFNVTVVRNDGDSVYTOVTFEPFLDLSYRKYSTLQONQSRQSWRLACSSASSTEV 840  
Db 796 GPREFNVTVVRNDGDSVYTOVTFEPFLDLSYRKYSTLQONQSRQSWRLACSSASSTEV 855

QY 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
Db 856 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 915

QY 901 QLELPVKYAVYVMVTSHGVTXKYNLFTASENTSRVMOHQVSNLQORSPISLIVLPV 960  
Db 916 QLELPVKYAVYVMVTSHGVTXKYNLFTASENTSRVMOHQVSNLQORSPISLIVLPV 975

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1020  
Db 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISAVCQRIQCDIP 1035

QY 1021 PFGIQEBFNATLKGSLFDMYIKTSHNHLIVSTABILFNDVSTFLLPGQAGFVRSQET 1080

Db 1036 PFGIQEBFNATLKGSLFDMYIKTSHNHLIVSTABILFNDVSTFLLPGQAGFVRSQET 1095

QY 1081 KVEFEFVENPLPLVGVSSVGLLALITAAALYKLGFFKROYKDMWSEGGPPGAEPQ 1137  
Db 1096 KVEFEFVENPLPLVGVSSVGLLALITAAALYKLGFFKROYKDMWSEGGPPGAEPQ 1152

RESULT 14  
US-08-476-062A-44  
; Sequence 44, Application US/08476062A  
; Patent No. 5877275  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY  
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,062A  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/216,081  
; FILING DATE: 21-MAR-1994  
; APPLICATION NUMBER: 07/637,830  
; FILING DATE: 04-JAN-1991  
; APPLICATION NUMBER: 07/539,842  
; FILING DATE: 18-JUN-1990  
; APPLICATION NUMBER: 07/212,573  
; FILING DATE: 28-JUN-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00786/068003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1163 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-476-062A-44  
Query Match 58.6%; Score 3446; DB 2; Length 1163;  
Best Local Similarity 60.5%; Pred. No. 5,1e-278;  
Matches 683; Conservative 143; Mismatches 297; Indels 6; Gaps 4;

QY 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVGAPEIVAANQSGSLYQCDYSTGSCPEI 60  
Db 20 FNLDTBELTAFRVDSAGFSGSVVQYANSWVVGAPQKITAANTGTGLYQCGYSTGACEPI 79

QY 61 RLQVPEAVNMSLGLSLAATTSPPQLLAGCPVHTQCSNTYVVKGLCLFLGSLNLRQOQPK 120  
Db 80 GLQVPEAVNMSLGLSLASTTSQLLACGPTVHEGGRNMYLTGLCFLGPT--OLTQR 137

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFFRMKELVSTIMEOLKSKTFLSLMOYSEEF 180  
Db 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFPATMNFVRAVISQFQRPSTQFSLMQFSNKF 197

QY 181 RIHFTFKEFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFTNGARKNAFKILFLL 240



541	QY	HGTSGSIGSPHSQR	IAGSKLS	SPRLQYFGQ	SLSGQDL	TMGLD	VLTVG	AGHVLLRSQ	600																																													
557	Db	HGVLGPS	:SPSHSQR	IAGSQL	SSRLQYFGQ	ALSGQDL	TQDGL	VDLAVG	ARGQVLLLR	616																																												
601	QY	PVLRVKAL	MEFNPRE	VARNV	PCNDQV	VKGK	EAGEVR	YCLHV	QKSTRDL	REGQIQSVVT	660																																											
617	Db	PVLVGVSM	QCFIT	PABIP	PRSAF	ECREQV	SEQTL	VQSNIC	LYIDK	SKNLLGSR	DQSSVT	676																																										
661	QY	YDLALD	SGRPHS	RAVEN	ETKNS	TRQTQ	VLGLT	QTCET	LKQL	PCNIED	PDVSP	IVLRNF	720																																									
677	Db	LDLALD	PGLS	PRATF	QETK	NSLS	SRV	VLGK	HA	CENFN	LLP	PCSV	EDSVT	PI	TRLNF	736																																						
721	QY	SLVGT	PLSA	FGNLR	PVLA	EDAQR	LFTAL	PF	FKNG	CNDNI	CO	DDLS	ITFS	PMS	LD	CLVVG	780																																					
737	Db	TLVGK	PLLA	FRNLR	PM	LALA	QRYFT	ASL	PF	FKNG	CADH	ICOD	NLG	ISF	SP	GL	KS	LV	796																																			
781	QY	GPREF	NVTY	VNR	DGED	S	YRTQ	VTFF	PLD	LS	VRKY	STLON	ORS	SR	ML	AC	ES	AS	TEV	840																																		
797	Db	SNLE	NAE	VW	NDG	SD	TG	TTT	F	SH	PAG	LS	YR	VA	E	Q	K	O	G	OL	RS	LH	T	C	D	S	APV	G	--	854																								
841	QY	SGALK	STCS	GIN	HI	P	ENSE	SV	TN	IT	F	D	V	D	S	K	A	S	L	G	N	K	L	L	K	A	N	T	S	S	N	N	P	R	N	K	T	E	F	900														
855	Db	SQGT	W	ST	C	S	R	I	N	H	L	I	F	R	G	A	Q	I	T	F	L	A	T	D	V	S	P	K	A	V	L	G	R	L	L	L	T	A	N	V	S	S	E	N	N	T	P	R	T	S	K	T	F	914
901	QY	QLE	P	K	V	K	A	V	M	V	M	T	G	H	G	V	S	T	K	L	N	T	A	S	-	E	N	T	S	R	V	M	O	H	Q	V	S	N	L	G	O	R	S	L	P	I	S	L	V	E	L	P	959	
915	Db	QLE	P	K	V	K	A	V	T	V	S	S																																										

Search completed: November 9, 2004, 12:20:59  
Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds  
(without alignments)  
4014.622 Million cell updates/sec

Title: US-09-902-481B-4  
Perfect score: 5884  
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMVSGGPPGAEPPQ 1137

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1 RWHULB	cell surface glyco
2	4460	75.8	1153	2 S00551	leukocyte surface
3	3469	59.0	1153	1 RWHULC	cell surface glyco
4	1543.5	26.2	1170	2 S03308	cell surface glyco
5	1527.5	26.0	1163	2 I56126	lymphocyte fuction
6	1140	19.4	1179	2 A53213	integrin alpha-E c
7	1093.5	18.6	1151	2 A45226	integrin alpha-1 c
8	1071	18.2	1170	2 I45314	integrin alpha-1 c
9	1057	18.0	1178	2 S44142	VLA-2 protein homo
10	1055	17.9	1180	2 A35854	integrin alpha-1 c
11	1054	17.9	1181	2 A33998	integrin alpha-2 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	642	10.9	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.4	1041	2 T31437	alphan integrin -
16	579.5	9.8	1034	2 JC7294	cell surface glyco
17	572.5	9.7	1051	2 A35761	VLA-3 alpha subuni
18	567.5	9.6	1053	2 I55534	integrin alpha-5 c
19	555.5	9.4	1053	2 S44250	integrin alpha-v c
20	546.5	9.3	1034	2 A36108	integrin alpha-6 c
21	535	9.1	1044	2 T10050	fibronectin recept
22	534.5	9.1	1072	2 A38457	integrin alpha-6 c
23	532	9.0	1049	2 A27079	integrin alpha-3 c
24	532	9.0	1073	2 B36429	integrin alpha-5 c
25	530.5	9.0	1051	2 A40021	integrin alpha-6 c
26	526	8.9	1048	2 A27421	integrin alpha-8 c
27	525.5	8.9	1091	2 A41543	integrin alpha-8 c
28	517	8.8	1044	2 I56516	integrin alpha-8 c
29	496.5	8.4	1146	2 S40311	integrin - fruit f

## ALIGNMENTS

## RESULT 1

## RWHULB

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein M.  
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004

C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; I52567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, C

B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:Cross-references: UNIPROT:P11215; GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:G307146

A:Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499,501-965,'P',967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA594

A:Note: the authors translated the codon TAC for residue 1129 as Thr

A:Note: part of this sequence, including the amino end of the mature protein, was confir

R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHB>

A:Cross-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:G553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhes

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor

A:Reference number: A32218; MUID:89098993; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HTC>

A:Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:G386975

integrin alpha-2b  
position-specific  
integrin alpha-7 c  
alpha-7 integrin -  
glycoprotein IIB -  
leukocyte adhesion  
integrin alpha cha  
F54P2.1 protein -  
integrin alpha v c  
hypothetical prote  
integrin alpha cha  
glycoprotein IIB -  
integrin alpha cha  
glycoprotein IIB -  
integrin alpha-1 -

30 495.5 8.4 1039 2 A34269  
31 494 8.4 1394 2 A29637  
32 489 8.3 1137 2 JCS950  
33 488 8.3 1135 2 I61186  
34 486.5 8.3 1037 2 A60163  
35 486 8.3 126 2 B30892  
36 470 8.0 1106 2 S38783  
37 462 7.9 1226 2 S44824  
38 453 7.7 1045 2 S60571  
39 444.5 7.6 1139 2 S28277  
40 426 7.2 1115 2 T09403  
41 421.5 7.2 1115 2 T09433  
42 391 6.6 764 2 I36916  
43 309.5 5.3 1086 2 T18523  
44 299 5.1 604 2 I36917  
45 286.5 4.9 272 2 A55348



R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986  
 A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep  
 A:Reference number: 159078; MUID:86287312; PMID:2942940  
 A:Accession: 159078  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 11-44 <RES>  
 A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193  
 C:Genetics:  
 A:Gene: Mac-1  
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo  
 C:Keywords: cell adhesion; glycoprotein; transmembrane protein  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental  
 F:148-318/Domain: von Willebrand factor type A repeat homology <WAA>  
 F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 75.8%; Score 4460; DB 2; Length 1153;  
 Best Local Similarity 73.8%; Pred. No. 2.4e-302;  
 Matches 840; Conservative 146; Mismatches 150; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFOENARGQGVVQLGSRVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 60  
 DB 17 FNLDTEHPMTFOENAKGFGQNVVLGGTSVVVAAPOEAKAVNQTGALYQCDYSTSRCHPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTPPQILLACGPTVHOTCSENTYKGLGCLFGLGSLNRQPOK 120  
 DB 77 PLQVPPEAVNMSLGLSLAVSTVPOQLLACGPTVHONCKENTYVNGLCYLFGLGSLNRPPQQ 136

QY 121 FPALRGCPQEDSDIAPLIDGSGSIIPHPFRMKEMWSTVMEQLKSKTFLSLMOXSEEF 180  
 DB 137 FPALRECPQESDIAPLIDGSGSINNIDFQMKFVSTVMEQFKSKTFLSLMQVSDPEF 196

QY 181 RIHFTKEFQNNPNSLIPITQLLGRTHATGLKRVRELFININGARKNAFKILFL 240  
 DB 197 RIHFTFDFNRNPSHSHVSPKQLNGRTKTSAGIRKVRRELFKHNGARENAKILVVI 256

QY 241 TDGEKFGDPLGYEDVTPELDREGVIRVIGVGAERSEKSRQELNTVASKPPRDHVFQIN 300  
 DB 257 TDGEKFGDPLDYKDVTPEDRAGVIRVIGVGNFKNPKQSRRELDTIASKPAGEHVFQVD 316

QY 301 NFALKTIONLREKIPIAIEGTQTGSSSFEHMSQEGFSAATISNGPLLSITVSGVSDWAG 360  
 DB 317 NFALNTIONQLEKIPIAIEGTQTGSSSFEHMSQEGFSAITISNGPLLSVGSFEDWAG 376

QY 361 GVFLYTSKEKSTFINTRVDSMDNDAYLGAAIILRNVRQSLVILGAPRYQHIGLVAMFR 420  
 DB 377 GAFLYTSKDKVTFINTRVDSMDNDAYLGASAVILNRVQSLVILGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYVEOTRGQSVCP 480  
 DB 437 ENFGTWEPHTSIKSQIGSYFGASLCSVDMDADGNTNLILIGAPHYVEKTRGQSVCP 496

QY 481 PRGQARWQCDVAVLGEQGPWRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 540  
 DB 497 PRG-RARWQCEALLHGDQHPWRFGAALTVLGVDNGDKLTDVAIGAPGEENRGAVYLF 555

QY 541 HCTSGSGISPSHSQRTAGSKLSPRLQYFQGSLSGGQDLTMDGLVLTVAQGHVLLRSQ 600  
 DB 556 YGASIASUSASHRIIGAHFSGLYFQGSLSGGKDLTMDGLMDLAVGAQGHLLLRQA 615

QY 601 PVLVRKAIMFNPREVARNVFCNDQVVKGEAVRVCLHVOKSTFRDLRGQTQSUTV 660  
 DB 616 PVLRLAETWTFPKKVARSVFACQEOVLKKNKQAGEVRVCLVRKNTKDLRGDTQSTVT 675

QY 661 YDLALDSGPHGRVAFNETKSTRQTQVLGHTQTCETLKLQPCIEDPVSPVILRLNF 720  
 DB 676 YDLALDPVRSIRAFDETKNTRRTQVGLMQRCETLKLILPCVDVDSVPIILRLNY 735

QY 721 SLVGTPLSAFGLRPVLAEDAQRLLFTALPPFKKNGDNICODDLSITFSFMSLDCLVVG 780  
 DB 736 TLVGEPLRSFGLRPVLAEDAQRFFTAFFPKKNGDNISICODDLSITWSAMGLDLTVVG 795

QY 781 GPRBNVTVVRNDGDSRYRTQVTFPFLDLSYRKVSTLQNRQSRL-ACESASSTE 839  
 DB 796 GPQDFNMSVTLRNDGDSYGTQVTVYVPSGLSYRKDSASQNLTKKFWFKPAESSSSSE 855

QY 840 VSGALKSTCSINHPPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMMPTNKTE 899  
 DB 856 GHGALKSTTNWNIHPIIPANSEVTFNITFDVDSHASFGNKLLKAIIVASENNMSRTHTK 915

QY 900 FQLELPVKYAVMVVTVTSHGVTSTKYNFTASENTSRVMQHOYQVSNLQORSPLISIVELVP 959  
 DB 916 FQLELPVKYALYMIYVTSDESSIRVNLNTASEMTSKVHQHQFNNGQSLVSVVFWIP 975

QY 960 VRLNQTVWRPQVTFNENLSSTCHTKERLPSHSDFLAELRKAPVAVNCIAVCORIQCDI 1019  
 DB 976 VQINNVTVDHPQVIFQSNLSSACHTEQKSPPHSNPRDQLERTFVLNCSVAVCKRIQCDL 1035

QY 1020 PFFGIQEEFNATLGNLSFDWYIKTNSHLLIIVSTAILNDSVFTLLPGGAFVRSQTE 1079  
 DB 1036 PSFNTQEIFNVTLKGNLSFDWYIKTNSHLLIIVSTAILNDSAFALLPGQESVRSKTE 1095

QY 1080 TKVEPEFVPNPLPIVGVSSVGGLLLLALITAAALYKLGFEKRYKDMSEGPPGAEPQ 1137  
 DB 1096 TKVEPEVBNPVLPIVGVSSIGGLVLLALITAGLYKLGFFRQYKDMNEAAPQADPPQ 1153

## RESULT 3

## RWHLIC

cell surface glycoprotein CD11c precursor - human

N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text\_change 09-Jul-2004

C:Accession: A36584; A35543; S00864

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 12750-12751, 1990

A:Reference number: A36584

A:Contents: erratum

A:Accession: A36584

A:Molecule type: DNA

A:Residues: 1-1163 <COR>

A:Cross-references: UNIPROT:P20702

A>Note: this revision to the sequence from reference A35543 includes the carboxyl end

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.

A:Reference number: A35543; MUID:90153906; PMID:2303426

A:Accession: A35543

A:Molecule type: DNA

A:Residues: 1-834 <CO2>

A>Note: this sequence has been revised in reference A36584

R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.

EMBO J. 6, 4023-4028, 1987

A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte

A:Reference number: S00864; MUID:88166645; PMID:3327687

A:Accession: S00864

A:Molecule type: mRNA

A:Residues: 1-755, 'L', 757-1163 <CO3>

A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830

A>Note: part of this sequence was confirmed by protein sequencing

C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my

C:Genetics:

A:Gene: GDB:ITGAX; CD11C

A:Cross-references: GDB:119758; OMIM:151510

A:Map position: 16p11.2-16p11.2

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homo

C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <SIG>

F:20-1107/Domain: extracellular #status predicted <EXT>

F:149-319/Domain: von Willebrand factor type A repeat homology <WAA>

F:1108-1133/Domain: transmembrane #status predicted <TM>

F:1134-1163/Domain: intracellular #status predicted <INT>

F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre



[illegible]

## RESULT 4

S03308

cell surface glycoprotein CD11a precursor - human

N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte

C;Species: Homo sapiens (man)

C;Date: 28-Feb-1990 #sequence,revision 28-Feb-1990 #text\_change 09-Jul-2004

C;Accession: S03308; A47458; A47565; A48759; S36044

C;Jarlsson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.

J. Cell Biol. 108, 703-712, 1989

A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha

A;Reference number: S03308; MUID:89139587; PMID:2537322

A;Accession: S03308

A;Molecule type: mRNA

A;Residues: 1-1170 <L>AR>

A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UBC8; EMBL:Y00796; NID:g31421; P

A;Note: part of this sequence was confirmed by protein sequencing

P;Cornwell, R.D.; Gollahon, K.A.; Hickstein, D.D.

Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993

A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or

A;Reference number: A47458; MUID:93248261; PMID:8097887

A;Accession: A47458

A;Molecule type: DNA

A;Residues: 1-20 <C>R>

A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBI:P:130863)

P;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993

A;Title: Identification of cell-specific and developmentally regulated nuclear

A;Reference number: A47565; MUID:93281759; PMID:8099450

A;Accession: A47565

A;Molecule type: DNA

A;Residues: 1-20 <S>HE>

A;Cross-references: GB:M95609

P;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.

J. Biol. Chem. 268, 19305-19311, 1993

A;Title: Characterization of the CD11a (alphaL, LFA-1alpha) integrin gene prom

A;Reference number: A48759; MUID:93374910; PMID:8103515

A;Accession: A48759

A;Molecule type: DNA

A;Residues: 1-20 <N>UE>

A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406

C;Genetics:

A;Gene: GDB:ITGAL; CD11A

A;Cross-references: GDB:119757; OMIM:153370

A;Map position: 16p11.2-16p11.2

C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A ri

C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface

F:1-25/Domain: signal sequence #status predicted <SIG>

F:164-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status p

F:164-1170/Domain: von willebrand factor type A repeat homology <VWA2>

Query Match 59.0%; Score 3469; DB 1; Length 1163;  
Best Local Similarity 60.9%; Pred. Mismatches 233; Indels 6; Gaps 4;  
Matches 688; Conservative 142;

1 ENLDTENAMTFOBNARGFGVSVQLQGSRRVVVGAPQBIIVAANORGSLYQCIDYTGSGCEPI 60  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
20 FNLDTIELTAFRVDVSAGFDSVVQYANSWVVVGAPQKITAAANTQTGLGYQCYSTGACEPI 79  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
61 RLQVPVEAVNMSLGLSAAATTSPPOLLAACPTVHQTCSENTYKGLCFELGSLNRQQPOK 120  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
80 GLQVPEAVNMSLGLSLASTTSPQLAACPTVHHCEGRNYLTGLCFLLGPT--QUTQR 137  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
121 FPBALRGCCQEDSDIAFLDNGSIIIPHDRFRMKEWSTVMEQLKKSKTLTFLSMOYESEF 180  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
138 LPVSRECPREQEODIVPLDNGSSISRNFAITMNFVRVISQPRFSTPFLMQFNKP 197  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
181 RIHTFKPEFQNPNRPSLIKEIPTQLLRGTHATGLRKVURBELFNITNGARKNAFKILPLL 240  
198 QIHTFFEEFRSTSNPLSLASVHLQGFVTYTATAIQNVHRLPHASYGARRDATKILIVI 257  
241 TDGEKFGDPDYEDVIPELDREGVIRVIYGVGDAPRSEKSRQELNTVASKPPRHVFQIN 300  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
258 TDGKKGDSLDYKDVI PWADAAGIIRYAIGVLAFQRNSWKELNDJASKPSQHIFAKE 317  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
301 NEFALKTIQNLREKIFAIECTQTGSSSPHEMSQEGFSAAITSNPGLLSTVGYSYWAG 360  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
318 DFDALDKDIQNLKEDI FAIBGTETTSSSFELEMAQEGFSAVFTPDGPVLGA VGSFTWSG 377  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
361 GVFLYTSKEKSTFINMTRVSDMDNDAYLGYAAA IILNRVOSLVIGAPPYOHIGLUVAMFR 420  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
378 GARLYPENMSPTFINMSQENVDMRDSYLGYSTELALMGVQSLLVLAGPYOHTGKAVIFT 437  
421 ONTWGMSNANVKGTQIGAYPGASLCSDVDNSGSTDVLVLIGAPHYYEQTRGGQVSVCLP 480  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
438 QVSRQWRMKAEVTTQTGSIYPGASLCSVDVTDGSTDVLVLIGAPHYYEQTRGGQVSVCLP 497  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
481 PRGORARWCDAVLVYGEOQPWGPFGAALTVLVDGVNGDKLDTAIVAGPEEDNRCGAVLYF 540  
498 PRGWR-RWWCDAVLYGEOCHPWGRFGAALTVLVDGVNGDKLTDVVIAGPEEENRCGAVLYF 556  
541 HGTSGSIGSPSHSQRIAGSKLSPRILOYFGQSLSGQDLTMOGLVDLTVGAQGHVILLRSQ 600  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
557 HGVLGPSISPSSHORSIAGSQLSSRLQYFGQALSQGQDLTQDGLVDLAVGARQVILLRT 616  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
601 PVLAKVAIMEPNPREVAENVPCNDQVYKGEAGVRVCLHVOKSTRDLRLEGQIQSVVT 660  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
617 PVLWVGWSMQPTPABTPRSAFECEQQVSEQVTLVQSNICLYIDKESKNLLGSEDLOQSVT 676  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
661 YDLALDGRPHSAVFNETKSTRQTQVGLGTQTCETLKQLPNCIEDPVSPITVLRUNF 720  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
677 LDIALDPCRLSPRAFTQETKNSLSRVRVLGKAHCENFNILLPSCVEDSVPTITRLNF 736  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
721 SLVGTPLSAPGNLRPVLBAEDAQRLETALPPPKNGCNONI CODDISITPESMSLDCLVVG 780  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
737 TLVGKLPLAFLNRLEMLAADAQRYFTASLPFFKNCADHICDNIGISFPGLCKSLVG 796  
||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : : ||||| : : :  
781 GPREFNVTVTVRNDEGDSYRTQVTEFFFLDLSYRKVSTLQNGSORSMWLACESASTEV 840  
797 SNLELNAEVWNWDEGDSYGTITTSHPAGLSYRVYAEQKQGLRSLHLTCDSPVGV-- 854  
841 SGALKSTSCSINHPIPPENSEVENITPFDVDSKASLGKLLKANVTSENMPRTNKITEF 900  
855 SQGTWSTSCKRIINHLIPRGAQITFLATFDVSPKAVLGDRLLLTANVSSNNTPRTSKTTF 914  
901 QLELPAVKAVYMVMTSHGVSTKYLNFTAS-ENTSRVMQHOYOVSNLGQBSLPSLFLVLP 959  
915 QLELPAVKAVYTVVSSHEQFTKYNLSBEEKESHVAMERYQVNNLQORDLPVSNFWWP 974  
960 VRINQTVIWDRPQVTFSENLSSCTCHKERLP SHSDFLAELRKAPVNVCISIAVCORIQCDI 1019  
975 VFIQNFAMVMDVEVSHPONPSLCSEKELAPPASDFLAHIQKNPVDCTAGCLRPFCDV 1034



```

QY 768 TFSFMSLDCLVGGP-----REFNVTVTVVRNDEGDSYRTQVTFPPPLDLISYRKVSTLQN 821
Db 780 SSPARS-----GFLRMSSASLAVETWTLNSGEDAYWVRLDLDFPRGLSFRKVEMLQ- 831
QY 822 QRSORSWRLACESASSTEVSGAL-KSTSCSINHPIPFENSEVTFNITFDVDSKASLGNKL 880
Db 832 --PHSRMPVSCSEL--TEGSSLTKTKCNVSSPIFKAGOEVSQVWFENILNLSWEDFV 887
QY 881 LLKANVTSEN-NMPTNKTEFQELPVKYAYVMVTVTSHGVSTKYKYNFTASENTSRVMQHQ 939
Db 888 ELNCTVHCENSSLOEDNSAATHIPVLPVNIILTKQENSTLYISFTPKGPKTQOVQHV 947
QY 940 YQVSNLQORSPLISLVLVPLVRLNQTVIMDRPQ-----VTFSENLS--TCHTKE-RLP 990
Db 948 YQV-----RIQPSAYDHNMPT-LEALVGPVRPHSEDLITYTWSQTDPLVTCHSEDLKRP 1001
QY 991 SHSDFELAEALRPVNVCSIAVCORIQCDIPFGIOEERFNATLKNGLSPDWIKTSHNHL 1050
Db 1002 SSE---ABQPCPLPGV-----QPRCPFIV---RWEILIQVTGIVELSKIKAS-STLS 1046
QY 1051 IVSTAELLFNDVSFTLLPGQAFVRSQTEVETKVEPEVENPILPLIVGSSVGGLLALLITA 1110
Db 1047 LCSSLVSFNSKPHLYGSKA-SEAQVLVKVDLHERKMLHVVYLVSGIGLVLLFLIFL 1105
QY 1111 ALYKLGFFKRYQKMM-SEGPPGAEP 1136
Db 1106 ALYKVGFFKRLKEXMEADGGVNGSP 1132

RESULT 6
A:Species: Homo sapiens (man)
C:Accession: A53213
R:Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha(E) subunit.
A:Reference number: A53213; PMID:94164962; PMID:8119947
A:Accession: A53213
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: UNIPROT:P38570; GB:L25851; NID:9457244; PID:9457245
A:Gene: GDB:ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.4%; Score 1140; DB 2; Length 1179;
Best Local Similarity 28.9%; Pred. No. 8.6e-71;
Matches 340; Conservative 214; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSLYQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATSPPPOLLACGTVHQ 95
Db 65 GPLHRCSLVQDEILCHPVEHVPIPKGRHGVTVVRSHHGLICI-----QVLVRRP--HS 117
QY 96 TCSNTYVKGCLFLGSLNRQQOQ----- 119
Db 118 LSSELT---GTCSLGGLDLPQAOANFFDLENLLDPARVDVTDGVCYNKGGGDDVNTA 174
QY 120 KFPALRGCPQED-----SDIAFLIDGSGIIPDFRRMKWVSTVMEQI-- 164
Db 175 RQRALKEEEDKEEEDDEEAEAGTEIALIDGSGIDPPDFQRAKDFSNMWRNRYE 234
QY 165 KKSRTLSLMQYSBEFRIHFTTFKQFNPNRSLIKPITQLLGHHTATGLRKVVRLEFN 224
Db 235 KCFECNFALVQGGVIQTEFLDRDSQVMSLARVQNTITGVSVTKTASAMQHVLDISFT 294
QY 225 ITNGARKNAFKILLFLITDGEFGDPLGYEDVIPELDREGVIRYVIGVGDAPFRSEKSRQEL 284

```

```

Db 295 SSHGSRKASKVMVLTLDGGIFEDPLNLTTVINSPKMQGVVERFAIVGVEFKSARTAREL 354
QY 285 NTVASPPRPHVFOINNFEALKTIONOLREKIFAIEGTQTGSSSSSEHEMSQBFSAIT 344
Db 355 NLIASPDETHAPKVTINMALDGLLSKLRIINISMEGT---VGDALHYQLAQIGFSAQL 411
QY 345 SNGP-LLSTVSGSDWAGGVFLY-TSKEKSTFINNTRVDSMDNA-----YLYGAAAIILRN 398
Db 412 DERQVLLGAVGDFWMSGALLYDTRSRGRFLNQTAAAAADAAQAQSYLYGYAVAVLHKT 471
QY 399 RVOSLVGAPRYQHI GLVAMFR-ONTGMWESNANV-KGTQIGAYFGASLCSVDVDSNGST 456
Db 472 CSLSYAGAPQYKHG--AVFELQKEGREASFPLVLEGEQMGSTFGSELCPVDIDMGST 529
QY 457 DLVLIGAPHYEQYTRGQVSCVPLRQGRARWQCDVLYGEGQGPWGRFGAALTVLGDNV 516
Db 530 DFLVLAAPFHVHGEERVVYRLSE-QDGSFSLARILSGHPGFTNARFEGFMAAAGDLS 588
QY 517 GDKLTDVAIGAP-----GEEDNR--GAVYFHGTSGSGLISPHSQRISAGLSPLRLOYFGQ 570
Db 589 QDKLTDVAIGAPLEGFGAGDGCASFSGSVIYNG-HWDGSLASPSQIRASTVAPGLQYFGM 647
QY 571 SLSGGQDLTMDGLVDLTGVAQCHVLLLRSQPLVAKATMEFNPREVARNVFECNDQVVKG 630
Db 648 SMAGGFDISGDGLADITVGTILGAVVFRSRVWRVKVSMATFSPALP-----1GF 697
QY 631 KEAGEVRVCLHVQKSTRDLREGQIQSVTVTDLALDSCRPHSRVAVNFETKSTRQTVL 690
Db 698 NGVNVRLCFEI-SSVTTASEGLREALNLTLDVGVKRRRLQCSQVRSCLGCLREWS 756
QY 691 GLTQTCETLKLQLPN-----CIEDPVSPVLRLNFSLVGTPLSAFGLNRPVLAEDARLFT 746
Db 757 SSSQLCEDL-LLMPTGELCEEDCFSNASVSYQL-QTPEGQTDHPQILDRTYEPFPAI 814
QY 747 ALFPEKKGNDNIQDDLSITFSMSLDCLVGGPRENVTVTVVRNDEGDSYRTQVTF 806
Db 815 FOLPYEKAKCNKLCFVAELOLA-TTVSQOELVVGTLKELTLNLTNSGDSYMTSMALN 873
QY 807 FPLDLSYRKVSTLQORSQRWELACESASTEVSGALKSTSCSINHPIPFENSEVTFNI 866
Db 874 YPRNIQ-----LKMOKPSPSNIQDDDPQV---ASVLMNCRIGHVPL-KRSSAHVS 923
QY 867 TFDVDSKASLGNKLLKANVTSENN---MPTNKTEFO---LELPVKAVVMVTVSHGV 919
Db 924 VQOLENAPNRTADITVTVTNSNRRSLANETHTLQERHGFVAVLSKPSIMVNTGQGL 983
QY 920 S---TKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVLVPLVRLNQTVIMDRPQVTFSE 977
Db 984 SHHKEFLPHVHGEN---LFGAEYQ-----LQICVPTKRLGLQVAAVKLRTQ 1028
QY 978 NLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIPFFGIGQEEFNATLKNLS 1037
Db 1029 ASTVCTWSEACAYSS-VQHVVEHWSVSCVIA-----SDKENVTVAEIS 1073
QY 1038 FQWYIKTSHNHLIVST-----AEILFNDVSFTLLPGQAFVRSQTEYKVEFEVNPPL 1091
Db 1074 WD-----HSEELLKDVTELQILGEISFNKSLGLENAENH---RTKITVVVLKDEKYHSL 1125
QY 1092 PLTVGSSVGGLLALLITALYKLGFFKRYQKMMSE 1128
Db 1126 PIIKSGVGGLLLVILVILFKGFFKRYQKQNLLE 1162

RESULT 7
A:Species: Homo sapiens (man)
C:Accession: A45226
R:Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 269, 2989-2996, 1993
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit

```

A:Reference number: A45226; MUID:93155124; PMID:8428973

A:Accession: A45226

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1151 <BRI>

A:Cross-references: UNIPROT:P56199

A:Experimental source: hepatoblastoma cell line HepG2

A:Note: sequence extracted from NCBI backbone (NCBIP:124326)

F:142-317/Domain: von Willebrand factor type A repeat homology <WAL>

Query Match 18.6%; Score 1093.5; DB 2; Length 1151;  
Best Local Similarity 27.4%; Pred. No. 1.5e-67;  
Matches 345; Conservative 206; Mismatches 487; Indels 203; Gaps 44;

```
QY 1 FNLDTENAMTFQENARG-FGQSVVOL---OQSRVVVGAQOEIVAAQNRQSLQCCDYSTGS 56
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1 FNVVDKNSMTSGPVEDMEGYIVQYENEEGKWKVLIGSLVGPQKNRTGDVYKCPVGRGE 60
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 57 CEP-IRLQVPEA-----VNMSLGLSLAATSPQOLLACGPTVHQTCSNTYVKGIL 106
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 107 CFLFGSNLRQOPKEPEALRCQPDSDIAFLIDGSGSIIPHDFPRMKEWSTVMEQLKK 166
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 120 CSDVSPPTQVNSIAP--VQBCSTQ-LDIVIVLDGNSIYPWDS-----VTAFLNDLLK 170
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 167 -----SKTLFSLMOYSEEFRIHFTFKFQNNPNRSLIKPTITOLLGR--THTATGLRKVV 219
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 171 RMDIGPKOTGVIVQYGENVTHEFNLNKYSTEEVLVAKKIVQGGQITMALGTDTAR 230
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 220 RELFNITNGARKNAFKILFLTDGKFGDPLGYEDVIPELDRGVIRYVIGVGDAFR--- 276
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 231 KEAFTEARGARRGVKVMVITDGGESH-DNHRLLKVIQDCEDENIQRFESIALGSYNRGN 289
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 277 --SEKSRQELNTVASKPRDRHVQINNFEALQIONREKIFAIEGTQTGSSSFEHEM 334
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 290 LSTKFEVEIKSIASEPTEKHFNFVDELALVTIVKTLGERIFALEATADQSAASFEMEM 349
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 335 SQEGFSAITNGPLLSLVGSYDAGGVFLYTSKE-----KSTF-INNTRVDSMDNDAYL 388
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 350 SQTGFSAHYSQDWMLGAVGYDNGVTVMQKASQIIIPRNTTFNVESTKNEFL-ASYL 408
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 389 GYAAAIILNRVQSL-VLGAAPYQHIGLVAMFRONTGMWESNANVKGTOIGAYGASLCS 447
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 409 GYTNSATASGDVLYIAGQRYNHTGVLIYRMEDGNKILQTLGSEIQISYFGSILTT 468
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 448 VDSDNSGTDVLVIGAPHY-----YEQTR-GGQVSVCPPLPRGQARWQCD 491
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 469 TDIDKSDNTDILLVGAPMYMTEKEBQGVVYVYALNQTREYQMSLEPIKQTCSSRQHN 528
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 492 AVLYGEOQOPWG-RFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAVLYLHGTSGSISP 550
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 529 SCTTENKNEPCGARFGTAIAVKDLNLDFGNDIVIGAPLEDDHGAIVTIYHG-SGKTIRK 587
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 551 SHSORIAGSKLSPQLVFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSOPVLEVKAIM 610
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 588 EYAQRIPIGGDGKTLKFFQSIHGEMDGLDGLTDVTIGLGAALFWSRDVAVVKVTMN 647
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 611 FNPREVARNPECNQDVVKGKAG--EVRVCLHVQ-KSTRDRLEBQIQSVVTVDLALDS 667
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 648 FEPKNVIQKNCH---MEGETVCINATVCFEYKLSKEDTIYEADLQ----YRVTLDS 700
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 668 GRHSRAVFNET-----KNSTRQTOVLGTQTCETLKLQLPNCIEDPVSPIVRLNFS 721
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 701 LRQISRFSGTQERKQVORITVKSEC-----TKHSFYMLDKHDFODSVR---ITLDEN 752
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 722 LVGTPLSAFQNLRPVLAEDAQRLTALFPFEKNCNDNIQDDLSITFFSFLSDCLVWVG 781
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 753 LT-DPENG-----PVLDDSLPNSVHEYIPPAKDCGNKEKICISLSLHVATTEKDLIVRS 806
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 782 PRE-FNVTVTRVDGSDSYTQVTFPPFLDSLVRKVTSTLQNRQSRQSLACSSASTEV 840
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 807 QNDKFNLSLTVKTKOSAYNTRTIVHYSPNLVFSGIEAIQKD-----SCSEN----- 853
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

#### RESULT 8

145914

Integrin alpha 2 subunit - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004

C:Accession: 145914

R:Kamata, T.; Fuzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A:Title: Identification of putative ligand binding sites within the I-domain of integrin

A:Reference number: A54402; MUID:94193647; PMID:7511592

A:Accession: 145914

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1170 <KAM>

A:Cross-references: UNIPROT:P53710; GB:L25886; NID:g439695; PIDN:AAB59255.1; PID:g43969

F:161-336/Domain: von Willebrand factor type A repeat homology <WMA>

Query Match 18.2%; Score 1071; DB 2; Length 1170;  
Best Local Similarity 27.4%; Pred. No. 5.5e-66;  
Matches 333; Conservative 219; Mismatches 495; Indels 168; Gaps 47;

```
QY 1 FNLDTENAMTFQ-ENARGFGQSVVOL---OQSRVVVGAQOEIVAAQNRQSLQCCDYST 54
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 19 YNVGLPKAKIPSGSPSESGFGYAVQOFTNPKGNMLLVGSPGFWPKNRMGDYKCPVDLST 78
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 55 GSCEPIRLQ-----VPVEAVNMNLSGLSLAATSPQOLLACGPTVHQTCSNTYVKGILC 107
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 79 TTCEKLNQITSTMSNVTEKTNMSLGLTLTRNVGTGGFLTCGFLPAQCGSQYTTTGV 138
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 108 FLFGSNLRQOPKEPEALRCQPDSDIAFLIDGSGSIIPHDFPRMKEWSTVMEQLK-- 165
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 139 SDVSPDF-QLRTSFAPAVQTCF-SFIDVVVVCDENSIYEPWD--AVKNFLEKFEVQGLDIG 194
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 166 KSKTFLSMOYSEEFRIHFTFKFQNNPNRSLIKPTITOLL-----GRTHATGLRKVVRE 221
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 195 PTQKQGLQYANNPRVFNLTNFKSKD---EMIKATQTFQYGGDLTNTFKAQIYARDT 251
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 222 LFNITNGARKNAFKILFLTDGKFGDPLGYEDVIPELDRGVIRYVIGV-----GDAFR 276
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 252 AYSTAAGRGPAKVMVVVTDGESH-DGSKLKAVIDQCKNDNIRFGIAVLGYLNRLALD 310
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 277 SEKSRQELNTVASKPRDRHVQINNFEALQIONREKIFAIEGTQTGSSSFEHEMSQ 336
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 311 TKLKIKEIKAIASIPTEHFHFNVSDADLEKAGATTIGEIQIFSIEGTVGQ-GDNFQMEMSQ 369
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 337 EGFSAAIT--SNGPLLSLTVSGVDWAGGVFLYTSKEKSTFINMT--RVDSDMN-DAYLGYA 391
Db |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

370 VGFSAEYSPQNNILMLGAVGAYDMSGTVVQKTPHGLHIFSKOAEQIILQDRNHSYLGY 429  
392 AAIILNRVQSLVLGAPRYOHIGLVAMPFRONTGMWESNANY-----KGTOIGAYFGASL 445  
430 VASISTGNSVHFVAGAPRANTVQGVLYSVN-----ENGNTVVIQSQRGDQIGSYFGSVL 484  
446 CSVDVDSNGSTDLVLIGAPHYYEQTR--GGQSVSCPLPRGORARWQCDVAVLYGEGQDPWG 503  
485 CAVDVKNDTITDVLVAGAPMYNDLKKBEGRVYLETITKG-ILNWH--QFLEGGNLENA 541  
504 RFGAALTVLGDNVNGKLTDAIAGPGEEDNRGAVLPHGTSGSGISPSHSORIASG--KL 561  
542 RFGSIAALSIDINMGDFNDVIVGSPLEQNSGAVIYNGHEGM-IRLAYSOKILGSDRAF 600  
562 SPRLOYGSGGGQDLTMDGLVDLTVGAQGHVLLRSQPVLRVKAIMEFNPREFVARNVF 621  
601 SSHLQYFGRSLDGYDGLNGDSITDVSAGFAFGVQVQLWSQSIADVSVDASFTEPKKI--TLL 658  
622 ECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVTYDLALD-----SGRPHSRVFN 677  
659 NKNABE-----KULKCF-----SAKFRPTQNNQVAIVYNTIDEDQESSRVISRGLEK 707  
678 ETKNSTRTQTVLGLTQCE--TLKQLPNCIEDPVSPVILRLNFSL--VGTPLSAFGNL 733  
708 ENNERCLOKTMIVSOAQRCEYIIHQEPS---DIISPLNLCWNISLENPGT-----756  
734 RPLVABDAQRLFTALFPFKKNGCNDNICODLSITP-----SPMSLDCLVVGPREFNVTV 789  
757 NPALAEYSETVTVFSIPPHKDCGDDGVCISDLVNVQQLPATQQQPFVIVSNQNKLTFSV 816  
790 TVRNDEGDSYRTQVTFPFFPLDLSYRKVSTLQNRQSRWSRLACESASST-EVSGALKSTS 848  
817 QLNKKESAYNTEIIVDFSENLF-----ASWSPVDGTEVTCQIASSQKSVT 864  
849 CSINHPFPEENSEVTNITFDVDSKASLGKGLKLLKANVTSENNMPRTNKTETQLELPVKY 908  
865 CNVGYPAKSKQQTFTINFDFNLQ-NLQNOASISPRALSESQEEENMADSNVNLKLSILY 923  
909 AVYVMVTSIGVSTKYLNFTASENTSRVMOHQVSNLGR-----SLPISLVFLV 958  
924 DAEIHIT-RSTNINFEVSLDGNVSVV-HSFE--DIGPKFTFSIKVTTGVSVPVSNMA---976  
959 PVRLNVTWDRPQVTFSEN---LSSTCHTKE-----RLPSSHDFLAE- 998  
977 -----SVIHIHPQYTKDNPLMYLTCVHTDQAGDISCAEINPLKIGQTSSSVSFKSEN 1030  
999 LRKAPVNGSIAYCQRIQCDIPFGIOBEFNATLKNLSPDWYIKTSHNHLIVSTAEL- 1057  
1031 FRHIKELNCKFTASCNIMCWLRLQVKGEYFLNVSTRWNGTEFAASTFTQVQLTAAAEID 1090  
1058 LFNDSVFTL-----LPGQAFVRSQTETKVEPFE-VPNPLPLIVGSSVCGILLALLITA 1110  
1091 TYNPQIYVIBENTVTP-----LATMKPHEKVEVPTGVIVGSVIAGILLALLALVA 1140  
1111 ALYKLGFFKRYQKDM 1125  
1141 ILMKLGFFKRYKEM 1155

RESULT 9  
S44142  
VLA-2 protein homolog - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C;Accession: S44142  
R;Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze  
submitted to the EMBL Data Library, January 1994  
A;Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not  
A;Reference number: S44142  
A;Accession: S44142  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-1178 <EDE>

A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:G473098; PIDN:CAA82877.1; PID:G473  
F;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.0%; Score 1057; DB 2; Length 1178;  
Best Local Similarity 27.8%; Pred. No. 5.3e-65;  
Matches 340; Conservative 210; Mismatches 487; Indels 188; Gaps 44;  
QY 1 FNLDTEAMTFQ-ENARFGQGVVQL-----QGSRVVVVGAPOEIVAAANQSGSYQC--DYST 54  
DB 27 YNVGLPGAKIFGSPSEQFGYSVQQLTNPQGNLLVGSFSGFENRMDGVVYKPCVDLPT 86  
QY 55 GSCSPIRLQ-----VPVEAVNMSLGLSLAATSPOLLACGPTVHOTCSENTVVKGLC 107  
DB 87 ATCEKLNQNSASISNVTETIKTNMSLGLTLTRNPGTGGFLTCGPLWAHQCGNQYATGIC 146  
QY 108 FLFGSNLRFQQQ--KPEEARLRCPOEDSDIAFLIDGSGSIIPHPFRMKEMVSTVMBQL 164  
DB 147 ----SDVSPDFQLTSPFAVQACUSL--VDVVVCDENSIYP--WEAKNKLFLVKFVTGL 199  
QY 165 K--KSKTLFSLMQXSEEFRIHFTFKEFQNNPNRSLIKPIQLLQ-RTHTATGLKRVRE 221  
DB 200 DIGPKKTOVALLIQYANEPRIIFNLNDFETKEDMVQATSETROHGGDLTNTFRAIFARDY 259  
QY 222 LFNITNGARKNAFKILFLTGTGKEGDPGLGYEDVIPELDREGVIRYVIGV-----GDAFR 276  
DB 260 AYSQTSGRPGATKVMVWVTDGESH-DGSKLKTVIQQCNDDEILRFGIAVLGYLNRNALD 318  
QY 277 SEKSQELNTVASKPRDHFVFOINNFEALKTIONOLREKIPIAIEGTQTGSSSFHEMSQ 336  
DB 319 TKMLKEIKATIASTPTERYFFNVABEALKEKAGTLGEQIFSIETVQG-GDNFQWEMAQ 377  
QY 337 EGFS--AITSNGPLLSTVGSYDAGGVFLYTSKSKSTFINMT--RVSDMN-DAYLGVA 391  
DB 378 VGFSAADYAPQNDILMLGAVGAFDWSGTLVQETSHKPVIPPKQAFQVQLDRNHSPLGY 437  
QY 392 AAIILNRVQSLVLGAPRYOHIGLVAMPFRONTGMWESNANV-----KGTOIGAYFGASLCS 447  
DB 438 VAAISTEDGVHFVAGAPRANTVQGVLYSVNK---QGNVTVIQSHRGDQIGSTFGSVLCS 494  
QY 448 VVDVDSNGSTDLVLIGAPHYYEQTR--GGQSVSCPLPRGORARWQCDVAVLYGEGQDPWG 505  
DB 495 VVDVDDKTTITDVLVAGAPTYMNDLKEEGKVLFTTKILNQHQ---FLEGPEGTGNARF 551  
QY 506 GAALTVLGDNVNGKLTDAIAGPGEEDNRGAVLPHGTSGSGISPSHSORIASKLSR- 564  
DB 552 GSAIAALSIDINMGDFNDVIVGSPVENENSGAVIYNGHQT-IRTKYSOKILGNSGAFRR 610  
QY 565 -LQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQPVLRVKAIMEENPREVARNVFEC 623  
DB 611 HLQFFGRSLDGYDGLNGDSITDVSIGALGVQLWSQSIADVAIEALFTP-----660  
QY 624 NQOVVKGKEAGEVRVCLHVQKSTRDLREGQIOSVVTYDLALD-----SGRPHSRVFN 679  
DB 661 -DKITLLNKDAKITLKLCFRASEFRPAGQNNQV--ALLFNMTLADAGHSRVTSGVFR 717  
QY 680 KNSTRQTVLGLTQTCET--LKLQPNCEIDPVSPVILRLNFSLIVGTPLSAFNGLRPVL 737  
DB 718 SERFLQNMVWNEVQKSEHHISIQKPS---DVVNPLDLRVDISLENPGTS-----PAL 768  
QY 738 AEDAQRLFTALFPPEKNCNDNICODLSI-----TFSFMSLDCLVVGPREFNV 788  
DB 769 EAYSETVKVFSIPFYKECGSDGICISDLILDVQQLPAIQQSF-----IVSNQKRLTFS 823  
QY 789 VTVRNDEGDSYRTQVTFPFFPLDLSYRKVSTLQNRQSRWSRLACESASST-EVSGALKST 847  
DB 824 VILKNGESAYNTVILAESENLF-----ASFSMPVDGTEVTCVSGSSQKSV 871  
QY 848 SC SINHPFPEENSEVTNITFDVDSKASLGKGLKLLKANVTSENNMPRTNKTET--FQLELP 905  
DB 872 TCDVGYPAKSKQQTFTINFDFNLQ-NLQNOAIAINFQAFSESQ--ETNKADNSVSTIP 928  
QY 906 VKIYAVMVVTSIGVSTKYLNFTASENTSRVMOHQVSNLGR-----SLPISLV 955

Db 929 LLYDAELHLLT-RSTWTFYIIGSDENAPSVIK---SVEDIGPKFIPSLKVTAGSPVMA 984  
 QY 956 FLV-----PVLNQTVIWDPPQVTF-SENLS 980  
 Db 985 LVTHIHPOYTKENPLLYLTGQTQOAGDISCTAEINPLKLPHTA-----PSVSFKNEFR 1040  
 QY 981 STCHTKERLPSPHSDFLAELRKAPVNVCSIAVQRIQCDDIPFFGIQIEEFNATILKGLSFDW 1040  
 Db 1041 ---HTKE-----LDCRTTSCNITCWLKDLHMAEYFINVTRVWART 1080  
 QY 1041 YIKTSHNHLIVSTAEILLPNDVSVFLLPQGGAFVRSQTEKVEPEVFNPLPIVSSVG 1100  
 Db 1081 FAASTFQVQTAAAEIDTHNPQLFVIEENAVTIPILMIMKPTKEAEVPT--GVLIGSIITA 1138  
 QY 1101 GILLALITALYALYKLGFFKQVYKDM 1125  
 Db 1139 GILLALLANTAGLWKLGFQKQYKDM 1163

RESULT 10  
 A35854  
 integrin alpha-1 chain precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 23-Oct-1990 #sequence\_revision 13-Sep-1991 #text\_change 09-Jul-2004  
 C/Accession: A35854; S11243  
 R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J.; Cell Biol. 111, 709-720, 1990  
 A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin A  
 A:Reference number: A35854; MUID:90338125; PMID:2380249  
 A:Accession: A35854  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-1180 <IGN>  
 A:Cross-references: UNIPROT:P18614; GB:X52140; NID:956493; PIDN:CAA36384.1; PID:956494  
 C:Keywords: cell adhesion; cytoskeleton; transmembrane protein  
 F:170-345/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 17.9%; Score 1055; DB 2; Length 1180;  
 Best Local Similarity 27.4%; Pred. No. 7,3e-65;  
 Matches 346; Conservative 196; Mismatches 475; Indels 246; Gaps 48;

QY 1 FNLDTENAMTQENARG-FGQSVVQL---QGSVVVVGAPQBIVAANQGSILYQCDYSTGS 56  
 Db 29 FNVDFKNSMSPGPFDMFGYTVQYENECKWVLIGSLPVGQPKARTGDYKCPVGRER 89  
 QY 57 CEP-IRLOQVPEA-----VMSLGLSLAATSPOLLACGPTVHOTCSENTYVYKGL 106  
 Db 89 AMPCVKLDLPNTSIPNVTEIKENNTFGSTL-VTNPNGGFLACGFLYAYRCGLHYTTGI 147  
 QY 107 CFLFGSNLRQOPKFPPEALRCPOEDSDIAFLIDGSGSIIIPHDFRMKEWSTVM---EQ 163  
 Db 148 CSDVSPTFQVNSFAP--VQECSTQ-LDIVIVLDGNSIYP-----NESVIAFLNDL 196  
 QY 164 LKK-----SKTLFSLMOYSSEFRTHFTKGFQNNPNRSLKIPITQLLG-RTHTATGLRK 217  
 Db 197 LKRMIDGPKQTQGVIGVQYGENVTNFEPLNLYKSYSTEEVLVAANKIGRQGLQTMALGIDT 256  
 QY 218 VVRELENTINGARKNAKILFLLDGKFGDPLGYEDVIELDREGVIRVIVGVGDAPR- 276  
 Db 257 ARKEAFTEARGARKGVKVMVIVTDGESH-DNVELKQVIOQCEDENIQRFSAIILGHYNR 315  
 QY 277 ----SKRSQELNTVASKPRDHFVQNNPFEALKTIQNLREKIFATEGTQSGSSSEFH 332  
 Db 316 GNLSTERFVEIKSIASEPTKEHFFNVSDDELALVTIVKALGERIFALENTADQSAASFEM 375  
 QY 333 EMSQEGFSAITSNGLPSTVGSYDWDAGGVFLYTSKEKSTFINMT--RVDSMDND---AY 387  
 Db 376 EMSQTGFSAHSQDWVMLGAVGYDNGVTVMQKQNMVPHNTTFTQTEPAKMEPLASY 435  
 QY 388 LGYAAAILLRVQSLVGLAPRQYHIGLVAMFRONTGWESNANVKQTQTQAYFGASLCS 447  
 Db 436 LGVTVNSATIPGDVLYTAGQFRYNTQVVIYKMEEDGNINILQTLGGEQIGSYFGSVLTT 495

QY 448 VVDVDSNGSTDILVLIGAPHY-----YEQTR-GQGVSVCPPLRQORARWQCD 491  
 Db 496 IDIDKSDSTDLLVGLAPMYMGTEKEEQGVVYAVVQNTREFYQMSLEPIRQTCSSLKDN 555  
 QY 492 AVLYGEOQOPWG-RFGAALTVLGDYNGDKLTDVAIGAPGEEDNRGAVLYLPHGTSGSGISP 550  
 Db 556 SCTKENKNEPCGARGTAIAVKDLNVDFNDVIGAPLEDDHAGAVIYIHG-SGKTIRE 614  
 QY 551 SHSQRIAGSKLSPRIQYFGQSLSGQDGLTMGDLVLTGGAQGHVLLLRSPQVLRKAIKE 610  
 Db 615 AVAQRIPSGGDKTKLFFGQSHGEMDLNGDLTDVTIGLGAALFWARDVAVVKVTNN 674  
 QY 611 ENPRVARNVPECNDQVVKGEAG--FVRVCLHVQ-KSTRDRRLRGQTSVVVYDLALDS 667  
 Db 675 FEPKNVNTQKXNCR---VEGKETVCINATMCFHVKLKSKEDSIYEDLIQ---YRVTLDS 727  
 QY 668 GRPSRAVFNET-----KNSTRQTOVLGLTQTCETLKLQLPNCI-----EDPVS 712  
 Db 728 LRQISRSPFGSGTQERKIQRNITVRESE-----CIRHSFYMLDKHDFQD 770  
 QY 713 PIVLRNLNFSLVGTPLSAFNGNLRPVLAEADAQRLFTALFFPEKNCGNDNICQDDLSITFSFM 772  
 Db 771 SVRVTLDENLTDPENG-----PVLDDALPNSVHEHIFPAKDCGNKERCISDLTLNVSTT 824  
 QY 773 SLDCLVYGGPRE-FNVTVTVNDEGDSYRTQVTPFFPLDLSYRKVSTIQNORSQSWRLA 831  
 Db 825 EKSLLIVKSHQDKFNVLSTVKNKGD SAYNTRTVVQHSNPLIFSGIEETQKD-----S 876  
 QY 832 CESASSTEVSGALKSTCSINHPIFFENSEVTEINTDVSQKASLGNKLL-LKANVTSEN 890  
 Db 877 CESN-----QNTICRVGYPFELRAGETVTFKLIIFQFNTSHLSENAIHLSTSDSEE 927  
 QY 891 NMPTNKTEFOLELPVKYAV-----YMWVTSHGVS-----KVLNFTASENTSRVMOHQ 939  
 Db 928 PLESINDNEVNIPIPVKEVGLQFYSSASEHHISVAANETIPEINST--EDIGNEINVF 985  
 QY 940 YQVSNLQORSILP---ISLVP-----LVFVRLNQTVIWD-----RP----- 971  
 Db 986 YTIRKRGHPMPQLQLSIFPNLTADGYPVLYPIG-----WSSSDNVNCRPSRLEDPFG 1039  
 QY 972 ----QVTF-----ENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQ 1016  
 Db 1040 INSKKMTISKSEVLKRGITQDCSSTC-----GVATITCSLLPSDSLQ 1082  
 QY 1017 CDI-----PFFGQIEEF---NATLKNLSFDWYIKTSHNHLIVSTAEILLPNDVSVFTLL 1067  
 Db 1083 VNVSLLLWKPTF-IRAHFFSLLNLTLRGELK-----SENSLTLSSN----- 1123  
 QY 1068 PQGAFVRSQTEKVEPEVFNPLPL--IVGSSVGGLLLLALITALYALYKLGFFKQVYKDM 1125  
 Db 1124 -----RKRELAIQISKDGLFGRVPLVILLSAFAGLLMLLILALWKIGFFKRPRLKK 1177  
 QY 1126 MSE 1128  
 Db 1178 MEK 1180

# RESULT 11 A33998

integrin alpha-2 chain precursor - human  
 N:Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 c1  
 C:Species: Homo sapiens (man)  
 C>Date: 30-Mar-1990 #sequence\_revision 18-Sep-1992 #text\_change 09-Jul-2004  
 C/Accession: A33998; B56793; A53117  
 R:Takada, Y.; Hemler, M.E.  
 J. Cell Biol. 109, 397-407, 1989  
 A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)  
 A:Reference number: A33998; MUID:89308879; PMID:2545729  
 A:Accession: A33998  
 A:Molecule type: mRNA  
 A:Residues: 1-1181 <TAK>  
 A:Cross-references: UNIPROT:P17301; GB:X17033; NID:933906; PIDN:CAA34894.1; PID:933907  
 A>Note: the authors translated the codon GAR for residue 802 as Gln, GTC for residue 803



R;Catimel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.  
Biochem. J. 279, 419-425, 1991  
A;Title: Separation of important new platelet glycoproteins (GP1a, GP1c, GP1c\*, GP1a and GP1b) from platelets of the human and mouse.  
A;Reference number: A56793; MUID:92061944; PMID:1953640  
A;Accession: B56793  
A;Molecule type: protein  
A;Residues: 30-43 <CAT>  
A;Experimental source: platelet  
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.  
J. Biol. Chem. 269, 463-469, 1994  
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements.  
A;Reference number: A53117; MUID:94103255; PMID:8276836  
A;Accession: A53117  
A;Molecule type: DNA  
A;Residues: 1-16, 'v', 18-21 <ZUT>  
A;Cross-references: GB:L24121; NID:9400342; PIDN:AAA16619.2; PID:94583535  
A;Note: authors translated the codon GTA for residue 17 as Leu  
C;Genetics:  
A;Gene: GDB:ITGA2; CD49B  
A;Cross-references: GDB:128031; OMIM:192974  
A;Map position: 5q11.1-5q11.2  
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein  
F;1-29/Domain: signal sequence #status predicted <SIG>  
F;30-1133/Domain: extracellular #status predicted <EXT>  
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>  
F;1134-1154/Domain: transmembrane #status predicted <TM>  
F;1155-1181/Domain: intracellular #status predicted <CY>  
F;1105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 17.9%; Score 1054; DB 2; Length 1181;  
Best Local Similarity 26.7%; Pred. No. 8.6e-65;  
Matches 327; Conservative 216; Mismatches 494; Indels 188; Gaps 43;

QY 1 FNLDTENAMTFQ-ENARGFQSVVQL-...QGSVVVGAQPEIVAAQNRSLYQC--DYST 54  
DB 30 YNVLPEAKIFSPSPSEQFGYVQVQFIPKGNWLLVGSPPENRMDGVYKCPVDLST 89  
QY 55 GSCEPIRLQ-----VPFEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTVYKGLC 107  
DB 90 ATCEKLNQTSIPNVTETKMTNLSGLILTRNMGTTGGFLTCQLPMAQQCGNYTTCVC 149  
QY 108 FLFGSNLRQOPQKPEALRCQPEDSDIAFLDGSIIIPHPRMKEWSTVMEQLK-- 165  
DB 150 SDISPPF-QLSASFSPATQCPSL-IDVVVVCDESIIYPWD--AVKNFLEKFPVQGLDIG 205  
QY 166 KSKTLFSLIMQYSEPRHFTTFKFFONNPNRSLIKPIITLLG-RTHATGLRKKVRELFN 224  
DB 206 PTKTQVGLIQYANNPRVFNLTNYKTEEMIVATSTQSYGGDLTNTFGAIOYARKYAS 265  
QY 225 ITNGARKNAKILFLITDGEKFGDPLGYEDVPELDREGVIRVIGV-----GDAFRSEK 279  
DB 266 AASGGRSATKVMVWVTDGESH-DGSKLKAVIDQCHNDNITRFGIAVLGVNRLNADTKN 324  
QY 280 SROELMTVASKPRDRHVQFNINFEALKTIONQRLKIFAIEGTQTGSSSSFEHMSQEGF 339  
DB 325 LIKEIKAIASIPERFFNVSDAALILEKAGTIGELIFSEGTIVQVQ-GDNFQEMSQVGF 383  
QY 340 SAATISNGP--LLSTVGSYDWAGGVFLYTSKESKSTFINMT--RVDSMN-DAYLGVAAAI 394  
DB 384 SADYSSQNDILMLGAVGAFGMSGTIVQKTSKSHLIPPKQAFDQTLQDRNHSYLGYSVAA 443  
QY 395 ILNRVQSLVGLAPRYOHIGVAMFRONTGWESNANV-----KGTQIGAYFGASLCSV 448  
DB 444 ISTGESTHFVAGAPRANTQGVLYSVN-----ENGNTIVQAHRGDQIGSYFGVLCV 498  
QY 449 DVDSNGSTDLVLGAPHYEQTR--GGQVSVCPPLRCORARQCDAVLVYEQGQWPGRFG 506  
DB 499 DVDKDTITDVLVCGAPMYMSDLKKEGRVYLFTIKKILGQHQ---FLEGPEGIENTRFG 555  
QY 507 AALTVLGVNGDKLITVAIGAPGEDNRGAVLPHGTSGGISPSHSORTAGS--KLSPR 564  
DB 556 SAIAALSDINMDGFNDIVGSPLENQNSGAVIYNGHQT-IRTKYSQKILGSDGAPRSH 614

565 LQYFQSGSLGGQDLTMDGLVLTVAQGHVLLLSQPVLRVKAIMEFNPREVARNVFCN 624  
DB 615 LQYFGRSLDGYGLNGDSITDVSIGAFQVQVQLWSQSIADVAIBASFTEPKI--TLVKNK 672  
QY 625 DQVVGKEAGEVRVCLHVQKSTRDLREGQIQSVVYDLDL-----SGRPHSAVFNETK 680  
DB 673 AQII-----LKLCF-----SAKPRPTKONNOVAIVNITLDAOGFSRVSRTSGLFGKNN 721  
QY 681 NSTRQTOVLGLTOPC--ETLKLQLPNCIEPVPSPVLRNLFNSILVGTSLPSAGNLRPVLA 738  
DB 722 ERCLOKQNMVQAQCPHIIYIQEPS--DVNSLDLRLVDSLENPGTS-----PALE 772  
QY 739 EDAQLFTALPFFKXNCGDNICQDLSITP-----SFMSLDCLVVGGRPFNVTVTRND 794  
DB 773 AYSETAKVFSIPFHKDCGEDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLKNK 832  
QY 795 GEDSVTRQVTFPPFLDLSYRKVSTLQORSQBSWPLACESAST-EVSGALKSTSCSINH 853  
DB 833 RESAYNTGIVVDFSENLF-----ASFSLPVDGTEVTQVAAQSKVACDVGY 880  
QY 854 PIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTEFOLELPKVAYVMV 913  
DB 881 PALKREQVTFITDFDNLQ-NLQNASLSPQALSQSEENKADNLVNLKIPLLYDAEI- 938  
QY 914 VTSHGVSTKYLNFATASENTSRVMQHQYQVSNIGQR-----SLPSLVFLV----- 958  
DB 939 ---HLTRSTNINFEISSDGNVPSIVHSPEVDGPKFIFSLKVTGTSVPVSMATVITHIQ 995  
QY 959 -----PVELNOTVIWDRPOVTF-SENLSSTCHTKER 988  
DB 996 YTKENPLMYLFGVQTDKAGDISCNADINPLKIGT-----SSVSFKSENFR---HTKE- 1047  
QY 989 LPSHSDFLAELKAPVWNCISIAVCORICDIPFFGIQEEFNATLKGNSLFDWIKYKTSNH 1048  
DB 1048 -----LNCRTASCNVTCWLKDXVHMKGEYFVNVTIRWNGTFASSTFOT 1091  
QY 1049 LLIVSTAETI-LFNDVSFTLLPQOGAFVRSGTETKVEPFVNP-----LP--LIVGSSVG 1100  
DB 1092 VQLTAAAEINTYNEIYVI-----EDNTVTIPLMIMKPDKAEVPTGTGIIGSIIA 1141  
QY 1101 GLLLLALITAAALYKLGFFKRYKDM 1125  
DB 1142 GILLLLALLVAILKLGFFKRYKDM 1166

RESULT 12  
A41131  
lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse  
N;Alternate names: integrin alpha-4  
C;Species: Mus musculus (house mouse)  
C;Date: 05-Jun-1992 #sequence\_revision 05-Jun-1992 #text\_change 09-Jul-2004  
C;Accession: A41131; S16742  
R;Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzmann, B.; Weissman, I.L.  
J. Cell Biol. 115, 1149-1158, 1991  
A;Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte  
A;Reference number: A41131; MUID:92064645; PMID:1840602  
A;Accession: A41131  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:951484; PIDN:CAA37316.1; PID:9514  
C;Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 663; DB 2; Length 1039;  
Best Local Similarity 22.7%; Pred. No. 1.3e-37;  
Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

QY 1 FNLDTENAMTFQ-ENARGFQSVV-QLQGSR--VVGGAPOEIVAAV-----QRGLYQCDY 52  
DB 41 YNLDPENALLYQGPSGLTFGYSVVLHSGSKRWLVIGAPTASWLSNAAVVPFGAIVRCGI 100  
QY 53 STG-----SCEPIRLQVP-----VEAVNMSLGLSLAATSTP-PQLLAG---PTVHQ 95



Db 101 RKNPNQTCQLQSGSPGPGCGKTCLEERDQWGLVTLRSRQGENGSIVTCGRWKNIFY 160  
QY 96 TCSNTYVYKGLCFGLSGNLROQPPPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKKE 155  
Db 161 MKSDNKLPTGICYVMPSDRLTSLK----- 185  
QY 156 MVSTVMEQLKSKTLFSLMOYSEEFRIHFTPEKQFNPNPRSLIKPITQLLGRTHATGL 215  
Db 186 -----RMAPCYADYT----- 195  
QY 216 RRVVRELFNITNGARKNAFKILFLITDGEKFGDPLGYEDVIPELDRGVIRYVIGWDAF 275  
Db 196 -----RKFG----- 200  
QY 276 RSEKSRQELNTVASKPRDHVFOINNFEALKTIONLREKIPAIEGTQTGSSSSFEHMS 335  
Db 201 -----NFAS-----C 205  
QY 336 QEGFSAITNSGPLLSTVGSYDWAGGVFLY--TSKEKSTFINMTVRVDSMDNDAYLGVA-- 391  
Db 206 QAGISSEFTQDLVWGAPGSSVWTGTVFYVYITNNQYKAFVD--RONQVKFGSYLGYSVG 263  
QY 392 AAILLRNVQSLVIGAPYQHIGLVAMFRQNTGMWESNAV---KGTQIGAYFGASLCS 447  
Db 264 AGHFRSPHTTEVVGAPQHEQIGKAYIF---SIDENELNIVYEMKKGKLSYFGASVCA 319  
QY 448 VDVDSNGSTDVLIGAPHYYQTRGGQVSVCPPLRGQRA--RWQCDVLYGQGGQPMWRFG 506  
Db 320 VDLNADGFSDL--LVGAPMQSTIREGRVFVY--INSGMGVAMVMEVRLVGLSDKYA--ARFG 376  
QY 507 AALTVLGVNDGDKLTDLVAIGAPGEDNRGAVYLVHGTSGSGISPSHSORLAGSKLSPLO 566  
Db 377 ESIANLGIDNDGFDIAIGAPQEDDLRGAVIYNGRV--DGISSTYSRIEQQOIKSLR 435  
QY 567 YFGSLSGQGLTMDGLVLDLVGA--QGHVLLRSQPLRVLRKAIEMFNPREVARNFECN 624  
Db 436 MFGQISGQIDADNNGYVDVAVGAFQSDSAVLLRTPWIVVEASLS--HPESVNTKFDCT 494  
QY 625 DVVYKGEAGEVRVCLHVQKSTRDLRGQIQSVVTVYDLDLDSGR---PHSRVAF--NET 679  
Db 495 -----ENGLPSVCMHLTLCFSYKGRVPGYIVLVNVDVHRKAESPRFYFNGT 547  
QY 680 KNSTRQTQVLGLTQTCETLKLQPLNCIEDPVSPIVLRNLSL-----VGTPLSAFGNLR 734  
Db 548 SDVITGSIYSSSEKCKTHQAFNRKQVDRLITFIHVEAYTHLGHVITKRNTEEFPLQ 607  
QY 735 PVLAEDQR--LFTALPFPEKNCNDNICODDLSTIFS-----FMSLDCLVVGPRBFNV 787  
Db 608 PILQKKEKDVIRKMINFARFCAVEN--CSADLQVSAKVGLPKPYENKTYLAVGSMKTI 666  
QY 788 TVTVRNDGDSYRQTVTFPFLDLSYKSVSTLQNRQSRWRLACESASSTEVSGALKST 847  
Db 667 NVSLFNAGDDAYETTLNVQLTGLYFIKILDEK-----QINCE-----VTESSGIVK-L 716  
QY 848 SCINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSEN--NMPRNTKTEFQLELP 905  
Db 717 ACSLGIYVDRLSRIDISFLIDVSSLSRAHEDLSIVHASCENECEGLDQVRDNRVLTIP 776  
QY 906 VKYAVVWVTSYGV--STKYINFTASNTSRVMOHQ-----YQVSNLQGRSLP--ISLVFL 957  
Db 777 LRYEV--MLTVHGLVNPSTSVYGSSENEPETCMAEKLNTLFHVINTGISMAPNVSKIM 834  
QY 958 VPVRLNQTIVDRP--QVTSSENLSSTCHTKE-----RLPSHSDFLAELR 1000  
Db 835 VP---NSFLPQDKFLNVLDVQTTTQGCFFKHYGRECTFAQKQGIAGTLTDIVKFLSKTD 891  
QY 1001 KAPVWNCIAA--VCQRIQCQDIPFEGIOEFENATLKGMLSDFWYIKTSHNLLIVSTAEIL 1058  
Db 892 KR--LLCYMKADQCHLDFLCN--FGKMEG-----KEASVHIQLEGRPSIL 933  
QY 1059 FNDVSFTLLPGQAFVRSQSTETKVPFEPVNP----- 1090

Db 934 EMDETSSL-----KFEIKATAPPEPHPKVIELNKDNVAHVFLGLHHQPKRHF 983  
QY 1091 -LPLIVGSSVGLLLALITLALYKLGFFKROYKDMWSE 1128  
Db 984 TIIITISLLGLLVILLISCVMMKAGFFKROYKSILOE 1022  
RESULT 13  
S06046  
integrin alpha-4 chain precursor - human  
N:Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain  
C:Species: Homo sapiens (man)  
C:Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text\_change 09-Jul-2004  
C:Accession: S06046; A39355; D28018  
R:Takada, Y.; Elices, M.J.; Crouse, C.; Hemler, M.E.  
EMBO J. 8, 1361-1368, 1989  
A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other inte  
A:Reference number: S06046; MUID:89356603; PMID:2788572  
A:Accession: S06046  
A:Molecule type: mRNA  
A:Residues: 1-1038 <TAK>  
A:Cross-references: UNIPROT:P13612; GB:X15356; EMBL:X15356; NID:g33945; PIDN:CAA34852.1  
R:Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.  
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991  
A:Title: Characterization of the alpha-4 integrin gene promoter.  
A:Reference number: A39355; MUID:91239513; PMID:2034655  
A:Accession: A39355  
A:Molecule type: DNA  
A:Residues: 1-93 <ROS>  
A:Cross-references: GB:M62841  
R:Takada, Y.; Strominger, J.L.; Hemler, M.E.  
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987  
A:Title: The very late antigen family of heterodimers is part of a superfamily of molec  
A:Reference number: A94151; MUID:87204112; PMID:3033641  
A:Accession: D28018  
A:Molecule type: protein  
A:Residues: 40-50, 'E', 52-53 <TA2>  
C:Genetics:  
A:Gene: GDB:ITGA4; CD49D  
A:Cross-references: GDB:I128032; OMIM:192975  
A:Map position: 2q31-2q32  
C:Superfamily: integrin alpha-4 chain  
C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transm  
F:I-39/Domain: signal sequence #status predicted <SIG>  
F:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>  
Query Match 10.98; Score 642; DB 2; Length 1038;  
Best Local Similarity 23.88; Pred. No. 3.8e-36;  
Matches 293; Conservative 175; Mismatches 409; Indels 356; Gaps 53;  
QY 1 FNLDTENAMTFQ-ENARFGQSVV-QLQGSR--VVVGAPQEIIVAAV---QRGSLYQCDY 52  
Db 40 YNVDTESALLYQGFHNTLFGYSVVLHSHGANRWLLVGAPTNLANASVINPQAIYKRI 99  
QY 53 STG---SCEPIRLQVPEAVNMSLGLSLAATTFPQLLACGPTVHOTCSNTYVKGLCFL 109  
Db 100 GKNPGQTCQLQSGP-----NGEP-----KTCLEERDQWL--- 133  
QY 110 FGSNLQQQPKPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKESVSVIMEQLKSKT 169  
Db 134 -GVTLRQPGE-----NGSIVTCGHR---W-----KN 156  
QY 170 LFSLMQYSEEFRIHFTPEKQFNPNPRSLIKPITQLLGRTHATGLRVVRELFNITNGA 229  
Db 157 IFVI----- 160  
QY 230 RKNAPKIFLLTDGEKFGDPLGYEDVIPELDRGVIRYVIGWDAPFRSEKSRQELNTVAS 289  
Db 161 -KNENK-----LPTGGCYGVF-----PDL-----RTELSKRIA 187  
QY 290 KPRDHFVFOI--NNFEALKTIQNLREKIPAIEGTQTGSSSSFEHMSQGFSAITNSGP 348  
Db 188 PCYQDIYKKGNFAS-----CQAGISSFTKDLI 217



Search completed: November 9, 2004, 12:19:28  
Job time : 33.25 secs

	Query Match	10.4%;	Score 614.5;	DB 2;	Length 1041;
	Best Local Similarity	24.8%;	Pred. No. 3.1e-34;		
	Matches 236;	Conservative 155;	Mismatches 314;	Indels 245;	Gaps 41;
QY	336	QEGFSAALTSNGP--LLSTVGSYDNAGGVFLYTSKEKSTFINNTRVDS----	DMNDAYLG	389	
Db	181	QAGFSGIIFSNSALVMGAPGSGYLGQYIVQVSSLNRSV--VQATQESNTGYTSPDNGYRG	239		
QY	390	YAAAI--ILNRVQSLVLGAPRYQHI--GLVAMFRONTGMWESNANVKGTGTOIGAVEGASLC	446		
Db	240	YSLALGDFNGGVQDVYVVGTPRAESLMGLVALFDQNLQFN---QWNGTQIVAFYGSVTV	296		
QY	447	SVVDVDSNGSTDLVLGAPHYHEVQTRGGQVSVCLPRGQARWQCDVLYGEQ-----	498		
Db	297	VVDI--NNDTYDGLVGAPMYMDGPAIQ-----RWEGAVVYVLQNPVDPGPGA	343		
QY	499	-----GPGWGFGAALTVLGDVNGDKLTDAIGAPGEEDNRGAVYLFHGTSGS	546		
Db	344	SNRLSSTLLGGQIRSRFGLSIGDSNQGFNDVALGAPYEGDDAGAVIYHG--SAN	402		
QY	547	GISPSHSORTAGSKLS--PRLOYFGOSLGGQDLTMDGLVDLTVGAQ--GHEVLLRSQPVL	603		
Db	403	GLKSTPAQVLTPTSLGHSGITTFGFSLOGQDMKNKYPDLLVGABSAANTAVLIRTPVV	462		
QY	604	RVAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQTQSV-----	658		
Db	463	SLDATLNTPE-----IGINLENKTYE--LADGTWVTSFIAMT	497		
QY	659	-----VTVDLALDSG--RPHSRVAFNETKNSRRQTQVLGL--TOTCETKLQ	702		
Db	498	CFTYTGNYLEDHIDISYTVTVDSGLIIANRAMFVNDMNSIETKTRRLAVSTQFCDPLRAY	557		
QY	703	LPNCIEDPVPVILURLNFSLVGTPLSAFCN-----LRPVLAEDAQRLLTALFPPEK	753		
Db	558	VGNSIEDKLPVKVLYQVLD-----NDESRLOPHEILPDIIMATMSTQTKQVSIQN	609		
QY	754	NCGNNDNICQDDLSTFTFSMLDCLVGGPREFNVTVTVRNDGEDSYRTQVTFVFFPLDLSY	813		
Db	610	NCVN--NICIPDLDTVTV--PNLPNIVIGOTQELTLDVSLNRRGEDAFQSSLSVYVYPLGLOF	667		
QY	814	RKVVSTLQNRQSQRSWRLACESASTVEYSALKSTSCSINHIFPEN-----SEVTFNIT	867		
Db	668	VRL-----ERKANMDFSVTSCSDS-----LRIITCDTGNPMVGKNILEGLTLTFQVS	717		



OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds  
(without alignments)  
4783.919 Million cell updates/sec

Title: US-09-902-481B-4

Perfect score: 5884

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKMMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	5836.5	99.2	1152	1	ITAM_HUMAN	P11215 homo sapien
2	4460	75.8	1153	1	ITAM_MOUSE	P05555 mus musculus
3	4371	74.3	1151	2	Q9J130	Q9i130 rattus norv
4	3799.5	64.6	920	2	Q28984	Q28984 sus scrofa
5	3469	59.0	1163	1	ITAX_HUMAN	P20702 homo sapien
6	3401	57.8	1162	1	ITAD_HUMAN	Q13349 homo sapien
7	3319.5	56.4	1188	2	Q6KAS4	Q6kas4 mus musculus
8	3319.5	56.4	1188	2	BAD21383	Bad21383 mus muscu
9	3310.5	56.3	1169	1	ITAX_MOUSE	Q9qxb4 mus musculus
10	3236.5	55.0	1161	1	ITAD_RAT	Q9qyb7 rattus norv
11	1572	26.7	1165	1	ITAD_BOVIN	P61625 bos taurus
12	1547.5	26.3	1170	1	ITAL_HUMAN	P20701 homo sapien
13	1546	26.3	1166	2	Q6TYB8	Q6tyb8 bos taurus
14	1546	26.3	1166	2	AAQ90015	Aaq90015 bos tauru
15	1527.5	26.0	1163	1	ITAL_MOUSE	P24063 mus musculus
16	1523.5	25.9	1161	2	Q9WTV4	Q9wtv4 mus musculus
17	1513	25.7	1160	2	Q9R200	Q9r200 mus musculus
18	1395	23.7	1196	2	Q98TF1	Q98tf1 cyprinus ca
19	1350.5	23.0	1086	2	Q96HB1	Q96hb1 homo sapien
20	1344.5	22.9	1187	2	Q98TF0	Q98tf0 cyprinus ca
21	1269	21.6	927	2	Q8HZV0	Q8hzv0 bos taurus
22	1154.5	19.6	1167	2	O88340	O88340 rattus norv
23	1148.5	19.5	1167	1	ITAE_MOUSE	P60677 mus musculus
24	1140	19.4	1179	1	ITAE_HUMAN	P38570 homo sapien
25	1109	18.8	1167	2	O88341	O88341 rattus norv
26	1093.5	18.6	1151	1	ITAI_HUMAN	P56199 homo sapien
27	1084.5	18.4	1189	1	ITAH_HUMAN	Q9ukx5 homo sapien
28	1071	18.2	1170	1	ITAD_BOVIN	P53710 bos taurus
29	1057	18.0	1178	1	ITAZ_MOUSE	O62469 mus musculus
30	1057	18.0	1178	2	Q6PIC7	Q6pic7 mus musculus
31	1057	18.0	1178	2	AAH65139	Aah65139 mus muscu

32	1055	17.9	1180	1	ITAI_RAT	P18614 rattus norv
33	1054	17.9	1181	1	ITAI_HUMAN	P17301 homo sapien
34	1051.5	17.9	1167	1	ITAG_HUMAN	O75578 homo sapien
35	1043.5	17.7	1188	1	ITAH_MOUSE	P61622 mus musculus
36	1043.5	17.7	1188	2	O7TQC3	O7tcq3 mus musculus
37	1041	17.7	1171	2	O42094	O42094 gallus gall
38	1032	17.5	1038	2	Q8BS01	Q8bs01 mus musculus
39	1005	17.1	895	2	Q9WUF8	Q9wuf8 mus sp. itg
40	998.5	17.0	1160	2	Q8MKF4	Q8mkf4 felis silve
41	865	14.7	348	2	Q8TES5	Q8tes5 homo sapien
42	860	14.6	1332	2	Q9BPQ8	Q9bpq8 halocynthia
43	796	13.5	205	2	Q63001	Q63001 rattus norv
44	770.5	13.1	304	2	Q6PG66	Q6pg66 mus musculus
45	770.5	13.1	304	2	AAH57200	Aah57200 mus muscu

## ALIGNMENTS

## RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	Name=ITGAM; Synonyms=CR3A, CD11B;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RT	J. Biol. Chem. 263:12403-12411(1988).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol. chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			



[illegible]





```

QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMER 420
Db 377 GAFLYTSKDKVTFINTRVDSMDNDAYLGYASAVILNRVQSLVGLGAPRYQHIGLVAMER 436
QY 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLGAPHYEYQTRGGQSVCP 480
Db 437 ENFGTWEPHTSIKSGQISYFGASLCSVDMDADGNTNLILGAPHYEYKTRGGQSVCP 496
QY 481 PRGQARWOCDAVLVGEQOPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRG-RARWOCDAVLVGEQOPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAVYLF 555
QY 541 HTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLMGLVGLDVTGVAQGHVLLRSQ 600
Db 556 YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGGQDLMGLVGLDVAQGHVLLRSQ 615
QY 601 PVLVRKATMEFNPREVARNVFECDNVKKGAGEVRVCHLVQKSTRDLREGQIQSVVT 660
Db 616 PVLRLATMEFSPKXVARSVFACQVNLKNDAGEVRVCLRVKNTKDLREGDIOSTVT 675
QY 661 YDLALDSGRPHSPRAVENETKSTRQTOVLGLTQTCETILKQLPNCIEDPVPVIRLNF 720
Db 676 YDLALDPVRSIRAFFDTKNTKTRRTQVFLGMQKCEYLKGLPDCVDSDSVPIILRLNY 735
QY 721 SLVGTPLSAFNGLRPVLAEADQRLFTALPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 736 TLVGEPLRSFNGLRPVLAMDAQRFTALPPEKNCNDNICQDDLSITFSAMGLDCLVVG 795
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWL-ACESASSTE 839
Db 796 GPQDFNMSVTLRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWL-ACESASSTE 855
QY 840 VSGALKSTCSINPIPPENSEVFNTFDVDSKASLGKLLKLLKANVTSENNMRTNTE 899
Db 856 GHGALKSTWINPIPPANSEVFNTFDVDSKASLGKLLKLLKANVTSENNMRTNTE 915
QY 900 FQLELPVKYAVYVMTSHGVSTKYINFNTASENTSRVMOHQVSNLQSRSLPISLVLP 959
Db 916 FQLELPVKYAVYVMTSHGVSTKYINFNTASENTSRVMOHQVSNLQSRSLPISLVLP 975
QY 960 VRLNQTVMDPQVTFSENLSTCHTKERLPSHSDFLAELKPAVYVNCSTIACQRIQCDI 1019
Db 976 VQIKNVTVMDPQVTFSENLSTCHTKERLPSHSDFLAELKPAVYVNCSTIACQRIQCDI 1035
QY 1020 PFFGQIEBFNATLKNLSFDWIKTSHNHLIVSTAEILFENDSVETLLPGQAVRSQTE 1079
Db 1036 PSENFQIEFVTLKNLSFDWIKTSHNHLIVSTAEILFENDSVETLLPGQAVRSQTE 1095
QY 1080 TKVEPFEVNPPLVGVSSVGLLALLALITAAALYKLGFFKRYQKDMSEGGPPCAEPQ 1137
Db 1096 TKVEPFEVNPPLVGVSSVGLLALLALITAGLYKLGFFKRYQKDMSEGGPPCAEPQ 1153

```

## RESULT 3

```

QJUI30 PRELIMINARY; PRU; 1151 AA.
AC QJUI30;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
OS Integrin beta 2 alpha subunit.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RA Fathallah D.M. Sr., Zeria K. Jr.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AF268593; AAF81280.1; -.
DR HSSP; P11215; 1BHQ.

```

```

DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR GO; GO:0007229; P:integrin-mediated signaling pathway; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 2.
DR Pfam; PF03357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int alpha; 5.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

```

## Query Match

Best Local Similarity 74.3%; Score 4371; DB 2; Length 1151;

Matches 826; Conservative 150; Mismatches 159; Indels 2; Gaps 2;

```

QY 1 FNLDTENAMTQENARGFGQSVVOLQGSRVVVGAPQEIIVAAANORGSLYCCDYSTGSCPEI 60
Db 17 FNLDTENPMTQENARGFGQSVIQLGETRVVVAAPQEVKAVNQTCALYCCDYSTNRCDEI 76
QY 61 RLQVPEAVNMSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFGLFNSLRQPOK 120
Db 77 PLQVPEAVNMSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFGLFNSLRQPOK 136
QY 121 FPEALRCGPQEDSDIAFLIDGSGSIIIPHDFRMKEWSTVMEQLKSKTFLSLMOYSEEP 180
Db 137 FPEALRCGPQESNIAFLIDGSGSINTIDFQKMKFEVSTVMDQFKSKTFLSLMOYSEEF 196
QY 181 RIHFTPFKFNPNPRSLIKPIITOLLGRTHATGLRVVRELENTGARGNAKILFL 240
Db 197 RTHFTFDFENPDKSHVPIRQINORTKASIRKVRVRELFOKINGARDNAKILVVI 256
QY 241 TDGKFGDPIGYEDVPELDRGVIRVIGVDGAFRSEKSRQELNTVASKPPRPHVQIN 300
Db 257 TDGKFGDPIGYEDVPELDRGVIRVIGVDGAFRSEKSRQELNTVASKPPRPHVQIN 316
QY 301 NFEALKTQNLQEKIFAIETGTGSSSSSEHEMSQEGFSAAITNSGPLLSTVGSYDWAQ 360
Db 317 NFEALNTIRNOLQEKIFAIETGTGSSSSSEHEMSQEGFSAAITNSGPLLSTVGSYDWAQ 376
QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMER 420
Db 377 GAFLYTSKDKVTFINTRVDSMDNDAYLGYASAVILNRVQSLVGLGAPRYQHIGLVAMER 436
QY 421 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLGAPHYEYQTRGGQSVCP 480
Db 437 QNTGWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLGAPHYEYQTRGGQSVCP 496
QY 481 PRGQARWOCDAVLVGEQOPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAVYLF 540
Db 497 PRG-RARWOCDAVLVGEQOPWGRFGAALTVDVNGDKLTDVAIGAPGEDNRGAVYLF 555
QY 541 HTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLMGLVGLDVTGVAQGHVLLRSQ 600
Db 556 YGASIASLSASHSHRIIGAHFSPGLQYFGQSLSGGQDLMGLVGLDVAQGHVLLRSQ 615
QY 601 PVLVRKATMEFNPREVARNVFECDNVKKGAGEVRVCHLVQKSTRDLREGQIQSVVT 660
Db 616 PVLRLATMEFSPKXVARSVFACQVNLKNDAGEVRVCLRVKNTKDLREGDIOSTVT 675
QY 661 YDLALDSGRPHSPRAVENETKSTRQTOVLGLTQTCETILKQLPNCIEDPVPVIRLNF 720
Db 676 YDLALDPVRSIRAFFDTKNTKTRRTQVFLGMQKCEYLKGLPDCVDSDSVPIILRLNY 735
QY 721 SLVGTPLSAFNGLRPVLAEADQRLFTALPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 736 TLVGEPLRSFNGLRPVLAMDAQRFTALPPEKNCNDNICQDDLSITFSAMGLDCLVVG 795

```

```
QY 781 GPREFNVTVVRNDEGSYRTQVTFEPFLDLSYKSVSTLQNRQSGRWRLACESASSTEV 840
Db 796 DSRDFDVSTVRNDEGSYGTVCYPSGLSYKVSASQNFQSKPVRVIAE-PSSEB 854
QY 841 SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 900
Db 855 QGVLSKSTIWDINHPIFFANSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNTEF 914
QY 901 QLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQBSLPSLVLVLPV 960
Db 915 QLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQBSLPSLVLVLPV 974
QY 961 RLNQTVIWRDQVTFSENLSSCTHCKERLPSHSDFLAELRKAPVNCISIAVCQRIQCIP 1020
Db 975 QINKVTIWDPPQVTFSENLSSCTHCKERLPSHSDFLAELRKAPVNCISIAVCQRIQCIP 1034
QY 1021 PFGIOEENFATLKGNSLSDWIKTSHNHLIVSTAEILFNDSVFTLLPQCGAFVRSOTET 1080
Db 1035 SPNSKEIFNVTLQGNLLFDWYIETSHDHLIVSTAEILFNDSVFTLLPQCGAFVRSOTET 1094
QY 1081 KVEPEVENPLPLVGVSSVGGILLALITAAIYKLGFFKQYKDMMSGGPPGABPO 1137
Db 1095 KVEPYTHNPVPLVGVSSVGGILLALITAGLYKLGFFKQYKDMMSGGPPGABPO 1151
```

## RESULT 4

```
Q28984 PRELIMINARY; PRT; 920 AA.
AC Q28984;
DT 01-NOV-1996 (TremblRel. 01, Created)
DT 01-FEB-1997 (TremblRel. 02, Last sequence update)
DT 01-JUN-2003 (TremblRel. 24, Last annotation update)
DE CD11b (fragment).
GN Name=CD11b;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; U40072; AAB16869.1; -
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS0234; VWF A; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON_TER 1
FT TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

```
Query Match 64.6%; Score 3799.5; DB 2; Length 920;
Best Local Similarity 78.8%; Pred. No. 3.2e-244;
Matches 726; Conservative 84; Mismatches 110; Indels 1; Gaps 1;
QY 118 PQKFPALRGCPQEDSIAFLIDGSGSIIPHDFRRMKVWSTVMEQKXSKTLFSLMQYS 177
Db 1 PQKFPALRGCPQEDSIAFLIDGSGSINRLDFRMKEFVSTVMGQFQKSKTLFALMQYS 60
QY 178 EEPRIHPTFEFQNNENPRSLIKPTQLGRTHATCLRKVPELENIUNGAKNAFKIL 237
Db 61 EDFYTHFTFNDKRNPSKLLVRPIQLGRTHATGIRKVRPELFHSGSGAKNAFKIL 120
```

```
QY 238 FLITDGEKFGDPLGYEDVIBELDREGVIRYVIGVDAPFRSEKSEQLNTVASKPRDHVF 297
Db 121 VITDGEKFGDPLGYEDVIBELDREGVIRYVIGVDAPFRSEKSEQLNTVASKPRDHVF 180
QY 298 QINNFEALKITONOLREKIFAIECTGTGSSSFFHEMSQEGFSAITNSGPLLSTVGSYD 357
Db 181 QVNNFEAVKTIQNLQKTFPAIEGTGTGSSSFFHEMSQEGFSAITNSGPLLSTVGSYD 240
QY 358 WAGGVLYTSKEKSTFNTMTVDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVA 417
Db 241 WAGGAFLEHMPKDRVIFNTTEVSDMDNDAYLGAAVEILNRQAQSLVLGAPRYQHTGLV 300
QY 418 MFRQNTGMWESNANVKGQIGAYFGASLCSVDVDSNGSTDLVLGAPHYYQTRGGQVSV 477
Db 301 MFKQNSGAWKKNADIKGSIQIGSYFGASLCSVDVNRDSSDLVLGAPHYYQTRGGQVSV 360
QY 478 CPLPRGQARQCDAVLYGEOGQWPGFQAALTVLGVNGDKLTDVAIGAPGEEDNRGAV 537
Db 361 CPLPQG-RAKWQCRVILGEOGHPWSRFGAALTALGVNGDKLTDVAIGAPGEEDNRGAV 419
QY 538 YLEHGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTLVGAQGHVLL 597
Db 420 YLEHGTSELGLSPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLMDLAVGAQGHVLL 479
QY 598 RSQPVLRVKAIMEFNPREVARNVPECNDQVVGKGEAGEVRVCLLVKQSTRDLREGQIQS 657
Db 480 RSQPVLRVEASVMVFKPREVARNVYECROAAKTQIAGEVQVCLQVRKSTWDRLEGGDTQS 539
QY 658 VVTYDLALDSGRPHSRVFNKTSRRTQVLGTQTCETLKLQLPNCIEDPVPVILR 717
Db 540 IITVDLALDPGRHPRAVFEETKNNTRRQTQTLGSRKCEHLALMLPDCVEDSVTPVILR 599
QY 718 LNFSLVGTPLGAFGNLRPVLAEADAQRLFTALFPFPEKNCNDNICQDDLSITFSMSLDC 777
Db 600 LNFSLVGPASSFGNLRPVLAVDAQRLFTALFPFPEKNCNDNICQDDLSITFSMSLDC 659
QY 778 VVGGRPRENVTVVRNDEGSYRTQVTFEPFLDLSYKSVSTLQNRQSGRWRLACESAS 837
Db 660 VVGGRDLKVTILTVRNQGEDSYRTQVTFEPFLDLSYKSVSTLQNRQSGRWRLACESAS 719
QY 838 TEVSGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNK 897
Db 720 TEESTALKSTSCSINHPIFFONSEVTNITFDVNDPDAFLGYKLLKANVTSENMPSSNK 779
QY 898 TEFQLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVMQHOYQVSNLQBSLPSLVL 957
Db 780 TEFQLELPVKYAVVMVTVSLEVTKYFNFTASEKTRHVEHQYQFNNLQQRKLPISVFEW 839
QY 958 VVRLNQTVIWRDQVTFSENLSSCTHCKERLPSHSDFLAELRKAPVNCISIAVCQRI 1017
Db 840 VVRLNRVTVMDPQVTFSENLSSCTHCKERLPSHSDFLAELRKAPVNCISIAVCQRI 899
QY 1018 DIPFGIOEENFATLKGNSL 1038
Db 900 DIPFGIOEENFATLKGNSL 920
```

## RESULT 5

```
ITAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20702; Q81VA6;
DT 01-FEB-1991 (Rel. 17, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN Name=ITAX; Synonyms=CD11c;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
```

RX MEDLINE=88166645; PubMed=3327687;  
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;  
 RT "cDNA cloning and complete primary structure of the alpha subunit of a  
 RT leukocyte adhesion glycoprotein, p150,95.";  
 RL EMO J. 6:4023-4028(1987).  
 RN  
 RP  
 RX MEDLINE=90153906; PubMed=2303426;  
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;  
 RT "Genomic structure of an integrin alpha subunit, the leukocyte p150,95  
 RT molecule.";  
 RL J. Biol. Chem. 265:2782-2788(1990).  
 RN  
 RP  
 RX MEDLINE=2388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McWay P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN  
 RP  
 RX MEDLINE=87167596; PubMed=3549901;  
 RA Miller L.J., Wiebe M., Springer T.A.;  
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1  
 RT and p150,95 leukocyte adhesion proteins.";  
 RL J. Immunol. 138:2381-2383(1987).  
 CC  
 CC -1- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It  
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell  
 CC interaction during inflammatory responses. It is especially  
 CC important in monocyte adhesion and chemotaxis.  
 CC  
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X  
 CC associates with beta-2.  
 CC  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC  
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in monocytes and  
 CC granulocytes.  
 CC  
 CC -1- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC  
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC  
 CC -1- SIMILARITY: Contains 1 VWFA domain.  
 CC  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11c.htm".  
 CC  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M81695; AAA51620.1; -.  
 DR EMBL; M29165; -; NOT ANNOTATED CDS.  
 DR EMBL; M29487; AAA51620.1; ALT SEQ.  
 DR EMBL; M29484; AAA51620.1; JOINED.  
 DR EMBL; M29483; AAA51620.1; JOINED.  
 DR EMBL; M29484; AAA51620.1; JOINED.  
 DR EMBL; M29485; AAA51620.1; JOINED.  
 DR EMBL; M29486; AAA51620.1; JOINED.  
 DR EMBL; BC038237; AAH38237.1; -.  
 DR PIR; A36584; RWHUIC.  
 DR PDB; 1N3Y; X-ray; A=141-338.  
 DR Genew; HGNC:6152; ITGAX.  
 DR MIM; 151510; -.  
 DR GO; GO:0008305; C:integrin complex; TAS.  
 DR GO; GO:0004872; F:receptor activity; TAS.  
 DR GO; GO:0007155; P:cell adhesion; TAS.  
 DR GO; GO:0009887; P:organogenesis; TAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWFA; 1.  
 DR PRINTS; PRO1185; INTEGRINA.  
 DR PRINTS; PRO0453; VWFADOMAIN.  
 DR SMART; SMO0191; Int alpha; 5.  
 DR SMART; SMO0327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 DR 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;  
 KW Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;  
 KW Signal; Transmembrane.  
 FT SIGNAL 1 19  
 FT CHAIN 20 1163 Integrin alpha-X.  
 FT DOMAIN 20 1107 Extracellular (Potential).  
 FT TRANSMEM 1108 1128 Potential.  
 FT DOMAIN 1129 1163 Cytoplasmic (Potential).  
 FT REPEAT 34 87 FG-GAP 1.  
 FT REPEAT 88 ? FG-GAP 2.  
 FT DOMAIN 165 351 VWFA.  
 FT REPEAT ? 401 FG-GAP 3.  
 FT REPEAT 402 453 FG-GAP 4.  
 FT REPEAT 455 517 FG-GAP 5.  
 FT REPEAT 518 576 FG-GAP 6.  
 FT REPEAT 581 633 FG-GAP 7.  
 FT CA\_BIND 466 474 Potential.  
 FT CA\_BIND 530 538 Potential.  
 FT CA\_BIND 593 601 Potential.  
 FT SITE 1131 1135 GPFKR motif.  
 FT DISULFID 69 76 By similarity.  
 FT DISULFID 108 126 By similarity.  
 FT DISULFID 655 712 By similarity.  
 FT DISULFID 771 777 By similarity.  
 FT DISULFID 848 863 By similarity.  
 FT DISULFID 998 1022 By similarity.  
 FT DISULFID 1027 1032 By similarity.  
 FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 392 392 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 697 697 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 735 735 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 899 899 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 939 939 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 1050 1050 N-linked (GlcNAc...) (Potential).  
 FT VARIANT 48 48 W -> R (in dbSNP:11574633).  
 FT FTID=VAR 018672.  
 FT CONFLICT 209 T -> S (in Ref. 4).  
 FT CONFLICT 251 T -> A (in Ref. 4).  
 FT CONFLICT 469 T -> S (in Ref. 4).  
 FT CONFLICT 490 G -> A (in Ref. 2).  
 FT CONFLICT 547 E -> K (in Ref. 4).  
 FT CONFLICT 756 D -> L (in Ref. 1).  
 FT CONFLICT 819 I -> V (in Ref. 4).  
 FT CONFLICT 1161 SEK -> TPHYQDNV (in Ref. 4).  
 FT STRAND 150 157



RL Gene 171:291-294 (1996).  
 RN [4]  
 RP INTERACTION WITH VCAM1.  
 RX MEDLINE=99059842; PubMed=9841932;  
 RA Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,  
 RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
 RT "alphabeta2 integrin is expressed on human eosinophils and functions  
 RT as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-  
 RT 1).";  
 RL J. Exp. Med. 188:2187-2191 (1998).  
 RN [5]  
 RP INTERACTION WITH VCAM1.  
 RX MEDLINE=99370002; PubMed=10438935;  
 RA Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
 RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
 RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
 RT binding interface between I domain and VCAM-1.";  
 RL J. Immunol. 163:1984-1990 (1999).  
 CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and  
 CC VCAM1. May play a role in the atherosclerotic process such as  
 CC clearing lipoproteins from plaques and in phagocytosis of blood-  
 CC borne pathogens, particulate matter, and senescent erythrocytes  
 CC from the blood.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D  
 CC associates with beta-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell  
 CC lines and subsets of peripheral blood leukocytes and strongly on  
 CC tissue-specialized cells, including macrophages foam cells within  
 CC atherosclerotic plaques, and on splenic red pulp macrophages.  
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; U37028; AAB38547.1; -  
 CC EMBL; U40274; AAB60634.1; -  
 CC EMBL; U40275; AAB60635.1; -  
 CC EMBL; U40276; AAB60636.1; -  
 CC EMBL; U40277; AAB60637.1; -  
 CC EMBL; U40278; AAB60638.1; -  
 CC EMBL; AF187881; AAF62875.1; -  
 CC HSSP; P11215; 1BHQ.  
 CC Genew; HGNC:6146; ITGAD.  
 CC MIM; 602453; -  
 CC GO; GO:0008305; C:integrin complex; TAS.  
 CC GO; GO:0016337; P:cell-cell adhesion; NAS.  
 CC GO; GO:0007160; P:cell-matrix adhesion; NAS.  
 CC GO; GO:0006955; P:immune response; NAS.  
 CC InterPro; IPR000413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP; 3.  
 CC Pfam; PF00357; Integrin\_alpha; 1.  
 CC Pfam; PF00092; VWA; 1.  
 CC PRINTS; PR01185; INTEGRINA.  
 CC PRINTS; PR00453; VWFADOMAIN.  
 CC SMART; SM00191; Int\_alpha; 5.  
 CC SMART; SM00327; VWA; 1.  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC PROSITE; PS00234; VWFA; 1.  
 CC Calcium; Cell adhesion; Integrin; Magnesium; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 17 Potential.

FT CHAIN 18 1162 Integrin alpha-D.  
 FT DOMAIN 18 1100 Extracellular (Potential).  
 FT TRANSMEM 1101 1121 Potential.  
 FT DOMAIN 1122 1162 Cytoplasmic (Potential).  
 FT REPEAT 32 85 FG-GAP 1.  
 FT REPEAT 86 ? FG-GAP 2.  
 FT DOMAIN 150 332 VWFA.  
 FT REPEAT 350 400 FG-GAP 3.  
 FT REPEAT 401 452 FG-GAP 4.  
 FT REPEAT 454 516 FG-GAP 5.  
 FT REPEAT 518 576 FG-GAP 6.  
 FT REPEAT 581 633 FG-GAP 7.  
 FT CA\_BIND 465 473 Potential.  
 FT CA\_BIND 530 538 Potential.  
 FT CA\_BIND 593 601 Potential.  
 FT SITE 1127 1131 GFEKR motif.  
 FT DISULFID 67 74 By similarity.  
 FT DISULFID 106 124 By similarity.  
 FT DISULFID 655 710 By similarity.  
 FT DISULFID 769 775 By similarity.  
 FT DISULFID 846 861 By similarity.  
 FT DISULFID 994 1018 By similarity.  
 FT DISULFID 1023 1028 By similarity.  
 FT CARBOHYD 59 59 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 87 87 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 99 99 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 391 391 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 691 691 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 733 733 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 873 873 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 957 957 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 1046 1046 N-linked (GlcNAc... ) (Potential).  
 FT CONFLICT 500 500 Missing (in Ref. 2).  
 FT CONFLICT 515 518 GHPW -> ATP (in Ref. 2).  
 FT CONFLICT 825 825 L -> V (in Ref. 2).  
 FT CONFLICT 984 984 V -> A (in Ref. 2).  
 SQ SEQUENCE 1162 AA; 126885 MW; F296A1A35455D77D CRC64;  
 Query Match 57.8%; Score 3401; DB 1; Length 1162;  
 Best Local Similarity 59.3%; Pred. No. 1.8e-217;  
 Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;  
 QY 1 FNLDTENAMTFQENARFGQSVVLQSGRVVGAPOEIVAAQNGSLYQCDYSTGSGCEPI 60  
 DB 18 FNLDVEEPTIFQEDAGFGQSVVQFGGSLRVVGAPELVVAAQNGSLYQCDAAATGMCQPI 77  
 QY 61 RLQVPVEAVNMSLGLAATTSPPOLLACGPTVHTCSENYVKGICFLFGSNLRQOPQK 120  
 DB 78 PLHIREAVNMSLGLTAASTNGSLRLACGPTLHRCVGENSYSGKSLGSRW-ETIQT 136  
 QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMKEWSTVMEOLKSKTLFSLMOYSEEF 180  
 DB 137 VPDATPECPHQEMDIVFLIDGSGSIDQDNFNQMGFQVAVMGQFEGTDLTFLALMOYSNL 196  
 QY 181 RHFTFKFQNNPNPSLTKPIQLLGRTHATGLKVVVRELNFITNGARKNAKILFL 240  
 DB 197 KHTFTFTQRTSPSQSLVDVPIQLKGLTFTATGILTVTQLFHHKNGARKAKKILVI 256  
 QY 241 TDGEKFGDPLGYEDVPELDRGVIRYVIGVDGDFRSEKSRQELNTVASKPRDHVQIN 300  
 DB 257 TDGQKYKDPLEYSDVIPAQEKAGIIRYAIGVGHAFQGPARTARQELNTISSAPPDHVKVD 316  
 QY 301 NFEALKTIQNLREKIFATEGTQTGSSSSFEHMSQEGFSAITSNGPLLTSTVSGVDWAG 360  
 DB 317 NFAALGSIQKLOEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFGLVGSFWSG 376  
 QY 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILRNRVQSLVGLGAPYQHGLVAMFR 420  
 DB 377 GAFLYPPNMSPTFINNSQENVDMRDSYLGYSTELALWKGVQNLVGLGAPYQHTGKAVFT 436  
 QY 421 QNTGWNESNANVKGTQIGAYFGASLCSDVDNSGSTDVLVIGAPHYEQTRGGQSVSCPL 480  
 DB 437 QVSRQWRKKAETVGTQIGSYFGASLCSDVDSDGSTDLLIGAPHYEQTRGGQSVSCPL 496

481 PRQARWQCDVAVLYGEQGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
497 PRQORVQWQCDVAVLYGRGQHWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556  
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 600  
557 HGASEGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 616  
601 PVLRVKALMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRLEGGIQSVVT 660  
617 PVLKVGVMREFSPVEKAVYRCWEKPSALEAGDAVCLTIQKSSLDQL-GDIQSVSR 674  
661 YDLALDGRPHSRVAFNETKXSTRQTVGLGTQTCETLKLQPLNCIEDPVSPIVRLNF 720  
675 FDLALDGRHLSTRAIPKTKTRALTRVKTGLNKHCHESVKLLLPACVEDSVPTLRLNF 734  
721 SLVGTPLSAFGLNLPVLAEDAQRFTALFPFEKNCNDNICODDLSITFFSMSLDCLVVG 780  
735 SLVREPISPQNLRLPVLAVSQDQLFTASLPFEKNCQDGLCEGLCVTLSPSGQLTVG 794  
781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSLACESASSTEV 840  
795 SSLELNVITVWVNNAGEDSYGTVSLYYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 853  
841 SGALKSTSCSNHPPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENNPRINKTER 900  
854 EG-LRSSRCVNHPIFHEGNSGTFTVTDVSYKATLGRMLMRASSENKASSKATF 912  
901 QLELPVKAVYVMTSHGVSUKYLNK-TASENTSRVMOHVOYVNSLQORSIPISVFLVP 959  
913 QLELPVKAVYVMTSHGVSUKYLNK-TASENTSRVMOHVOYVNSLQORSIPISVFLVP 972  
960 VELNQTIVLWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQDI 1019  
973 VLLNGVAVWVWVMEAPSQSL--PCVSEKPPQHSDFLTQISRSPMLDCSIADCLQPCDV 1030  
1020 PFFGIQEEFNATLKNLSPDWIKTSHNHLIVSTABILFNDSVFTLLPGQAFVRSQTE 1079  
1031 PSFSVQEEELDFTLKGNLSFGWRETLQKKVLVSVVAEITFDTSVVSQLPQGEAFMRAQME 1090  
1080 TKVEFEVFNPLPLVINGSSVGLLALALATLALYKLGFFKRYKQVNMWSE 1128  
1091 MVLEEDVYNAIPIMSSVGLLALALATLALYKLGFFKRYKQVNMWSE 1139

RESULT 7  
Q6KAS4  
ID Q6KAS4 PRELIMINARY; PRT; 1188 AA.  
AC Q6KAS4;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE MFLJ00114 protein (Fragment).  
GN Name=MFLJ00114;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCHI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,  
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;  
RA "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:  
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs  
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly  
RT Sampled from Size-Fractionated Libraries.";  
RL DNA Res. 11:167-180(2004).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein. (By similarity).  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
DR EMBL; AK131133; BAD21383.1; -  
DR GO; GO:0009897; C:external side of plasma membrane; IDA.  
DR InterPro; IPR000413; Integrin\_alpha.

InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_alpha; 1.  
DR Pfam; PF00092; VWA; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR0453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; UNKNOWN\_1.  
DR PROSITE; PS0234; VWEA; 1.  
KW Cell adhesion; Integrin; Transmembrane.  
FT NON TER 1  
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;  
Query Match 56.4%; Score 3319.5; DB 2; Length 1188;  
Best Local Similarity 56.8%; Pred. No. 4.9e-212;  
Matches 647; Conservative 171; Mismatches 303; Indels 19; Gaps 7;  
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAANORGSVLYQCDYSTGSCBPI 60  
DB 39 FNLDAEKPTEFHMDDGAEFGHSLVQYDSSVWVVGAPKEIKATNQIIGGLYKCGYHTGNCBPI 98  
QY 61 RLOVPVAVNMVSLGLSLAATTPPQLLACGPTVHTQTCSENTYVKGCLFGLFNSLRQQPQK 120  
DB 99 SLOVPPPEAVNMVSLGLSLAATNPWLLACGFTVHTCENIYLTGLCFLLSSTSPFKQS-QN 157  
QY 121 FPEALRGCPQSDSDIAFLIDGSGSIIIPHDPRRMKEWSTVMEQLKSKTSLFSLMOYSEF 180  
DB 158 FPTAQCEPKQDDIVFLIDGSGSISSTDFEKLMDFKAVMSQLQRPSTFSLMQFSDF 217  
QY 181 RIHFTFKFQNNPNRSLIKITQLLGRTHATGLRKVVRELENTGARKNAKILFL 240  
DB 218 RVHFTFNNFISTSSPLSLDLSVRQURGYTYTASAIKHVITELFTTQSGARQATKVLVI 277  
QY 241 TDGEKFGPLGVEDYVIBELDREGVIRYVIGVDGDAFRSKSRQELMTVASKPRDHVQFN 300  
DB 278 TDGRKQGNLSDVSDVPMVAEASIIRYAIGVSKAFNEHSKQELKAIASMPSEHYVFSVE 337  
QY 301 NPEALKTTQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNGPILLSVTGSDWAG 360  
DB 338 NFDALKDIEQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTPDGVLGAGVGSFWSG 397  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNAYLGAAAILLRNVSQSLVLCAPRYOHGLVAMPR 420  
DB 398 GAFLYPSNMRPTFINMSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTKVI 457  
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLIAPHYIEQTRGGQVSCPL 480  
DB 458 QESRHRWRPKSEVRGTQIGSYFGASLCSVDMDRDSGLDLVLIQVPHYIEHTRGGQVSCPM 517  
QY 481 PRQARWQCDVAVLYGEQGWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
DB 518 P-GVGRWHCGTTLHGEGQHPWGRFGAALTVLGDVNGDSLADVAIGAPGEENRGAVYIF 576  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQHVLLRSQ 600  
DB 577 HGASRQDIAPSPSQRIASQIPSRITQYFGQSLSGGQDLTRDGLVDLVLAGSKRVLLNTR 636  
QY 601 PVLRVKALMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDRLEGGIQSVVT 660  
DB 637 PIURVSPVTHFTFAEISRSVFECQVQVAPETLSDATVCLHIHESPKTL--GDLRSTVT 694  
QY 661 YDLALDGRPHSRVAFNETKXSTRQTVGLGTQTCETLKLQPLNCIEDPVSPIVRLNF 720  
DB 695 FDLALDGRHLSTRAIPKTKTRALTRVKTGLNKHCHESVKLLLPACVEDSVPTLRLNF 754  
QY 721 SLVGTPLSAFGLNLPVLAEDAQRFTALFPFEKNCNDNICODDLSITFFSMSLDCLVVG 780  
DB 755 SLVGVPISSLQNLQPMPLAVDDQTYFTASLPFEKNCNADHICQDDLSVVFGLDKTLVVG 814  
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPFLDLSYRKVSTLQ-----NQRQRQWR 829  
DB 815 SDLELNVDTVVSNDGDSYGTITVTLFYFVGLSPRVAEGQVFLRKCKEDQQRQGHSLH 874



Query Match	56.4%;	Score	3319.5;	DB 2;	Length	1188;			
Best Local Similarity	56.8%;	Pred. No.	4.9e-212;						
Matches	647;	Conservative	171;	Mismatches	303;	Indels	19;	Gaps	7;
QY	1	FNLDTENAMTFQENARGFCQSVVOIQGSRVVGAPQETVAANQROGSLYQCCDYSTGSCPEI	60						
Db	39	FNLDIAEKPTHFHMDDGAEGFHSVLQYDSSWVVVGAPKEIKATNIQIGLYKCYHTGNCPEI	98						
QY	61	RLQVPEAVNMSLGLSLAATTSPPOLLAGCPVTHQTCSENTYVVGKLCFLFQSNLRQPOQK	120						
Db	99	SLQVPEAVNMSLGLSLAATNPMSLLACGPTVHHCTRENITYLTGCLFLLSSSFQKS-QN	157						
QY	121	FPEALRGCPQEDSDTAFLIDGSGSIIPHDFRMKEWSTVMNEQLKKSKTFLSLMQYSEEF	180						
Db	158	FTAQOCEKQPODQVFLIDGSGSISSTDFEKLMDLFVKAVMSQLQRPSTFSLMQPSDYF	217						
QY	181	RLHFTFKCFQNNPNRSLLIKPTITQLLGRTHATGLRKVVRELFNITNGARKNAKPIFLFL	240						
Db	218	RVHFTFNFISSPLSLDVSRLQURGYTYSAAIKHVITELFTTQSGARQDATKVLIVI	277						
QY	241	TDGEKFGDPLGVEDVIPBLDREGVIRYVIGVDGAFRSEKSRQELNLTWASKPPRDHVFQIN	300						
Db	278	TDGRKQGNLSYDSVIPMAEASIIIRYAIGVKAFYNHSEKQELKATASMPSHSYFVSVE	337						
QY	301	NPEALKTTIONLREKIPALEGTQGTQSGSSSFHEMSQSGFSAAITNSGPLLSTVSGYDWAQ	360						

RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,  
RA Tsuchiya H.;  
RT "Isolation of genes selectively expressed by dendritic cells.";  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibronogen. It  
CC recognizes the sequence G-P-R in fibronogen. It mediates cell-cell  
CC interaction during inflammatory responses. It is especially  
CC important in monocyte adhesion and chemotaxis (By similarity).  
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X  
CC associates with beta-2 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
CC with I-domains do not undergo protease cleavage.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF211864; AAF23492.1; --  
DR HSSP; P20702; 1N3Y.  
DR MGD; MGI:96609; ltagx.  
DR GO; GO:0009897; C:external side of plasma membrane; IDA.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_alpha; 1.  
DR Pfam; PF00092; VWFA; 1.  
DR PRINTS; PRO1185; INTEGRINA.  
DR PRINTS; PRO0453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWFA; 1.  
DR PROSITE; PS00242; Integrin\_ALPHA; 1.  
DR PROSITE; PS00234; VWFA; 1.  
DR Calcium; Cell adhesion; Integrin; Magnesium; Receptor;  
KW Repeat; Signal; Transmembrane.  
FT SIGNAL 1 19 Potential.  
FT CHAIN 20 1169 Integrin alpha-X.  
FT DOMAIN 20 1116 Extracellular (Potential).  
FT TRANSMEM 1117 1137 Potential.  
FT DOMAIN 1138 1169 Cytoplasmic (Potential).  
FT REPEAT 34 87 FG-GAP 1.  
FT REPEAT 88 ? FG-GAP 2.  
FT REPEAT 152 330 VWFA.  
FT REPEAT ? 402 FG-GAP 3.  
FT REPEAT 403 454 FG-GAP 4.  
FT REPEAT 456 518 FG-GAP 5.  
FT REPEAT 519 577 FG-GAP 6.  
FT REPEAT 582 634 FG-GAP 7.  
FT CA\_BIND 467 475 Potential.  
FT CA\_BIND 531 539 Potential.  
FT CA\_BIND 594 602 Potential.  
FT SITE 1140 1144 GFFKR motif.  
FT DISULFID 69 76 By similarity.  
FT DISULFID 108 126 By similarity.  
FT DISULFID 656 711 By similarity.  
FT DISULFID 770 776 By similarity.  
FT DISULFID 858 873 By similarity.  
FT DISULFID 1007 1031 By similarity.  
FT DISULFID 1036 1041 By similarity.  
FT CARBOHYD 89 89 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 267 267 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 393 393 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 734 734 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 949 949 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 1059 1059 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 1084 1084 N-linked (GlcNAc... ) (Potential).

SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;  
Query Match 56.3%; Score 3310.5; DB 1; Length 1169;  
Best Local Similarity 56.6%; Pred.No. 1.9e-211;  
Matches 645; Conservative 173; Mismatches 303; Indels 19; Gaps 7;  
QY 1 FNLDENAMTFQENARGFGSGVVLQGSRRVVVVGAPQEIIVAAANRGSLYQCDYSTGSCBPI 60  
DB 20 FNLDAEKLTHFMDGAEFGHSLVQLYDSSVVVVGAPKEIKATQKQIGGLYKCYGHTGNCBPI 79  
QY 61 RLQVPVAVNMSLGLSLAATTPPQLLACGPTVHTCTSENTYVYKGLCFLFGSNLRQQPQK 120  
DB 80 SLQVPEPAVNSLGLSLAATTPSWLLACGPTVHTCTSENTYVYKGLCFLLSSSPKQS-QN 138  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMEKWSVTWEOQLKSKTFLSLMOYSEEF 180  
DB 139 PPTAQCEPCQDQDIVFLIDGSGSISSTDFEKLDFKAVMSLQRPSTFSLMOFSDYF 198  
QY 181 RIHFTFKFQNNPNRSLIKITQLLGHHTATGLRKVVRELFNITGARNNAKILFLL 240  
DB 199 RVHFTFNFISTSSPLSLGSRVQLRGVYTTASAIKHVITELFTTQSGARQDATKVLVI 258  
QY 241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPFSKSKQELNTVASKPRDHFVQIN 300  
DB 259 TDGRKQGNLSYDSVIPMAEASIIRYAIGVGKAFYNEHSKQELKAIASMPGSHYVFSVE 318  
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPEHEMSQEGFSAAITNGPLLSVTGSDWAG 360  
DB 319 NFDALKDIENQLKEKIFAIEGTETPSSSTFELEMSQEGFSAVFTDPGVLAGVGSFSWG 378  
QY 361 GYFLYTSKEKSTFINNTRVDSMDNAYLGYAAIILNRVQSLVLGAPYQHIGLVAMFR 420  
DB 379 GAFLYPSNMRTPTFINMSQENEDMRDAYLGYSTALAPFWKGVHSLILGAPRHQTKGVIFT 438  
QY 421 QNTGWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVLIGAPHYYEYOTRGQSVCPPL 480  
DB 439 QESRHRPKSEVRGTQIGYFGASLCSVDMDRDGSTDLLVIGVPHYYEYHTRGQSVCPM 498  
QY 481 PRGQARWQCDVLYGEOQOPWGRFGAALTGLVDVNGDKLTVAICAPGEEDNRGAVYLF 540  
DB 499 P-GVGRWHCGTTLHGEQHPWGRFGAALTGLVDVNGDSLADVAICAPGEENRGAVYIF 557  
QY 541 HGTSGSGISPSHQRIAGSLPRLQYFGOSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600  
DB 558 HGASRQDIAPSPQSIRISASQIPSRIQYFGOSLGGQDLTRDGLVDLVAVGSKRVLLLRTR 617  
QY 601 PVLVRKAIMEFNPREVARNVFCNDQVWKGKAGEVRVCLHVOKSTRDLRLEGIOQSVIT 660  
DB 618 PILRVSTVHTFTFAEISRSVFEQEQVAPQTLSDATVCLHIHESPKTL--GDLRSTVT 675  
QY 661 YDLALDSGRPHSAVFNETKNSSTRQTVGLTQCTETKLQLPNCIEDPSPVLRLNF 720  
DB 676 FDLALDHGLSTRATFAFKETKTRALTRVKTGLNKHCSVKLLLPACVEDSVPTITLRLNF 735  
QY 721 SLVGTPLSAFNLPPVLAEDAQRLLFTALFPKCKNGNDNICODDLSITFSFMSLDCLVVG 780  
DB 736 SLVGVPISSLQNLQPLAVDDQTYFTASLPKCKNGADHICQDLSVVFPGDLAKTLVVG 795  
QY 781 GPREFNVTVVRNDGEDSYRQTVTFPPFLDLSYKRVSTLQ-----NQSRQSWR 829  
DB 796 SDLELVNDVTVNSDGEDSYGTVTLFYPVGLSFRVABGQVFLRKKEDQWQRGQSHLH 855  
QY 830 LACESASSTEVSGALKSTSCSINHPIPEENSEVFNITFDVDSKASLGNKLLKANVTSE 889  
DB 856 LMCD--STPDRSQGLWSTSCSRHVFIRGGSQMTFLVTFDVSXPAELGDRLLRLARVGE 913  
QY 890 NNMPTNKTETQELPVPKAYVMVVTSEGVSTKYLNFTASE--NTSRVMQHQYQVSNLGR 948  
DB 914 NNVPCTPKTTFQELPVPKAYVMVVTSEGVSTKYLNFTASE--NTSRVMQHQYQVSNLGR 973  
QY 949 SIPISLVFLVPLVRLNQTIVWRPQVTFSENSLSTCHTKERLPSSHSDFLAELRKA PVNCS 1008  
DB 974 DVEFVSINFWPIELGEAVW--TVMVSHQPNLPTQCYVRNRLKPTQFDLLTHMQKSPVIDCS 1032

QY 1009 IAVCORTQDIPFGIOEFENATLKNLSFDWYIKTSHNHLIIVSTRAILLFNDVSFTLLP 1068  
D 1033 IADCLHRCIDPSILGDLDELFLKGNLSFGWISQTLQKKVLLSEAIETNTSVISQLP 1092  
QY 1069 GOGAFVRSQETKVEPEVENPLPLIVGSSVGGLLALITAAALYKLGFFKROKDMWSE 1128  
D 1093 GQEAFLRAQTKTLEMYKHNVPPLIVGSSVGGLLALITAILYKAGFFKROKEMLEE 1152

RESULT 10  
ITAD RAT  
ID ITAD RAT STANDARD; PRT; 1161 AA.  
AC OSQVE;  
DT 05-JUN-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Integrin alpha-D precursor.  
GN Name=Itgad;  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=101116;  
RN [1]  
RC SEQUENCE FROM N.A.  
RP STRAIN=Sprague-Dawley;  
RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,  
RA Gallatin W.M.;  
RT "Cloning of rat alpha D, a novel beta 2 integrin."  
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and  
CC VCAM1. May play a role in the atherosclerotic process such as  
CC clearing lipoproteins from plaques and in phagocytosis of blood-  
CC borne pathogens, particulate matter, and senescent erythrocytes  
CC from the blood (By similarity).  
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D  
CC associates with beta-2 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
CC with I-domains do not undergo protease cleavage.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- SIMILARITY: Contains 1 VWFA domain.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
DR EMBL; AF021334; AAF21241.1; --  
DR HSP; P11215; 1BHQ.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWFA\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_alpha; 1.  
DR Pfam; PF00092; VWFA; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWFA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PS0234; VWFA; 1.  
KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
Repeat; Signal; Transmembrane.  
FT SIGNAL 1 19 Potential.  
FT CHAIN 20 1161 Integrin alpha-D.  
FT DOMAIN 20 1100 Extracellular (Potential).  
FT TRANSMEM 1101 1121 Potential.  
FT DOMAIN 1122 1161 Cytoplasmic (Potential).  
FT REPEAT 34 87 FG-GAP 1.

FT REPEAT 88 ?  
FT DOMAIN 152 334 FG-GAP 2.  
FT REPEAT 352 402 VWFA.  
FT REPEAT 403 454 FG-GAP 3.  
FT REPEAT 456 517 FG-GAP 4.  
FT REPEAT 519 577 FG-GAP 5.  
FT REPEAT 582 634 FG-GAP 6.  
FT CA\_BIND 467 475 FG-GAP 7.  
FT CA\_BIND 531 539 Potential.  
FT CA\_BIND 594 602 Potential.  
FT SITE 1126 1130 GFPR motif.  
FT DISULFID 69 76 By similarity.  
FT DISULFID 108 126 By similarity.  
FT DISULFID 656 711 By similarity.  
FT DISULFID 769 775 By similarity.  
FT DISULFID 845 860 By similarity.  
FT DISULFID 993 1017 By similarity.  
FT DISULFID 1022 1027 By similarity.  
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).  
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).  
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.0%; Score 3236.5; DB 1; Length 1161;  
Best Local Similarity 57.5%; Pred. No. 1.6e-206;  
Matches 648; Conservative 163; Mismatches 303; Indels 13; Gaps 9;

QY 2 NLDTENAMTFQENARGFQSVVLOGSRVVGAPQEIIVAAQNGSLYQDYVSTGSCPIR 61  
D 21 NLDVEEPTVFREDAAAFQGTQVVGSGSLVVGAPLEAVAVNQTGLYDCAPATGMCPIV 80  
QY 62 LQVPVEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLRQOPKF 121  
D 81 LRSPLAEVNMISLGLSLVLTATNNAQLACPTAQRACVKNYAKGSCLLGSSL-QFIQAV 139  
QY 122 PEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKEMVSTVMEQLKKSKTLFSLMOYSEBFR 181  
D 140 PASMPECPQEDSDIAFLIDGSGSINQDFQKDFKALMGFEASTSLFSLMOYSNILK 199  
QY 182 IHFTFKFQNNPNRSLIKPTTOLLGTHATGKVVURELPTNIGARKNAFKILFLT 241  
D 200 THFTFKKILLDQSLVDPIVQOGUTYATGRTVMEELFHSKNGSRKSAKKILLVIT 259  
QY 242 DGEKFGDPLGYEDVPIPELDREGVIRYVIGVDGAFRSEKSKROELNVTASKPDRHVFQINN 301  
D 260 DGQYRDPLEYSDVIIPADKAGIIRYALVGDAFQEPALKELTIGSAPDQDHVFKVGN 319  
QY 302 FEALKTTONLRKIFAIEGTQSGSSFEHMSQEGFSAATSNGLPLSTVSGSYDWAG 361  
D 320 FAALRSIQRLQOEKIFAIEGTQSGSSFEHMSQEGFSAATSDGVLGVNGVFSWGG 379  
QY 362 VFLYTSKEKSTFTNMTRVDSMDNDAYLGAAAILNRNVSQSLVLCAPRYQHIGLVAMFRQ 421  
D 380 AFLYPNTRPTFTNMQENVDMDSDVILGYSTAVAFNKGVHSLILGAPRHQHTGVWFTQ 439  
QY 422 NTQMWSNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYVETQGGVSCVPLP 481  
D 440 EARHWRPKSEVRGTQIGSFAGSLCSVDVDRDGSSTDLVLGAPHYVETQGGVSCVFPVP 499  
QY 482 RQORARWQCDVLYGEGQPGWGFAGALTIVLGVNKGDKLTDVAIGAFGEEDNNGAVYLFH 541  
D 500 -GVRGEWQCEATLHGEGQHPWGRFVALTVLGVNKGDNLDVAIGAFGEESRGAVYIFH 558  
QY 542 GTSGSGISPSHSQRIAGSKSLRPLQFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSOP 601  
D 559 GASRLIMPSPQRVTGSQLSLRLQYFGQSLSGGQDLTQDGLVDLAVGAGQGHVLLRSIP 618



QY 239 LLTDGEKFGDPLGVEDVPELD-REGVIRVYVIGVDAPRSEKSRQELNTVASKPRDHFV 297  
Db 258 IITDGEATDE-----HNIDAADKDIRIIGIKGNFKTKESQEAHQFASKEVEFVK 309  
QY 298 QINNFEALKTIQONREKIFAIEGTQTQSSSPHEHMSQSFSAITNSNPLSLTVGSYD 357  
Db 310 ILDTFEKLDLFYELQKIVIEGTQKQDITSFNMELSSSGISADLSEGHVGVAVGAKD 369  
QY 358 WAGVGF-LYTSKSKSTFINNTRVDSMDNDAYLGAIA-IILNRVQSLVLGAPRYQHIGL 415  
Db 370 WAGGFLDLKADLKSSTVGNELTVESRAGLYGVTVLPSRGTMSLLATGAPRYQHVGR 429  
QY 416 VAMFQO--NFCMGENANVKGTOIGYFASLCSVDVDSNGSTDVLVLGAPHYEQTRGG 473  
Db 430 VLLFQDPKRGPSQIQEIDGIGISYFGGELCGVDVDRDGETELLIIAAPLYGQRG 489  
QY 474 QVSVCLPRQARWQCDVLYCEQGPWGRFGAALTVLGVNDGDKLTVAICAPGEEN 533  
Db 490 RVFIY--QKIQLEFQWSELQETGYPLGRFGAAIAALTIDNGDELTDVAVGAPLEE-- 544  
QY 534 RGVYLFHGTSGGISPSHSQRIAGSLSPRLQYFQCSLGGQDLTMDGLVDLTVGAQGH 593  
Db 545 QGAVYIFNGQOG-GLSPRPSQRIEGTQMPFGSGIQWFGRSHGVKDLGGDLADVAVGAGQ 603  
QY 594 VILLRSQPLRVKAIMENPREVARNVFCNDQVVKKEAG-EVRVCLHVQK--STRDLR 650  
Db 604 VIVLSRPVVDITTSVSPAPIPVHEVECSYSTSNQKKEGVNLTVCFOVKSLSIT---- 659  
QY 651 REGQISQVVTYDALDSGPRHAFVNETKNSRQTVGLTQICETLKLQPCIEDP 710  
Db 660 FQGHVLANITYLQDCHTRSGGLFPGGKHKLIGNTAVTPV-KSCFVFWFFHCIOQL 718  
QY 711 VSPVILRNFSL---VGTPLS--AFGNLRFVLAEDAQRLLFTALFFPKKNGCDNLCQDDL 765  
Db 719 ISPINVLSYSLWEIEGTPRDPRALDRDIPPLTKPSPLETKEIFEKNCGDKNCEADL 778  
QY 766 SITEFSMLDCLVCGPREFNVTVVRNDGDSYRQVTFPPFLDLSYRKVSTLQNSRQ 825  
Db 779 KLAFSMRKILRLTPSASLVRPLRNTAEDAYVWQVTLSPQGLSFRKVEIL---KPH 835  
QY 826 RSWRLACESASTVSGALKSTCSINHPIPENSEVTNITFDVDSKASLGNKLLKAN 885  
Db 836 SHVPVGCHELEPAVHS-RALSCNVSPFIPEGSDMVDIQMNTLQKSGWGFELQAN 894  
QY 886 VTS-----ENNPRNTKTEFQLEPLVKYAVVTVSHGVSKYLNFTASENTRVMQ 937  
Db 895 VSCNEDSSLEDNSATTS-----IPVYPINVLTKQENSTLYISFTPKSKIHVVK 947  
QY 938 HOYQV---SNLQORSPLISVLVPLVPLNQTVI---WD---RFQVTFPS-ENLSSTCHTK 986  
Db 948 HIYQVRIQPSNYDNP-PLEALVVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESDEA 1006  
QY 987 ERLPSSHDFLAELKAPVWNCIAVCQRIQCDIPFGIOEFNATLKNLSFDWYKTS 1046  
Db 1007 E-----SCSPGT--EFCPIDF---RQELVQVNGVWELRGTIKAS- 1042  
QY 1047 NLLIVSTAELFNDSVFTLLPGQAFVRSQETETKVEPFEVNPPLVIGSVSGVGLLA 1106  
Db 1043 SMLSCLSLAISNSKHFLYGRNASM-AQVWKVLDVVEKEMLYLVLSGIGGLLLP 1101  
QY 1107 LITAAVLKLGFKHQYKDM-----SEGGPPGAPQ 1137  
Db 1102 LIFALYKGVFFKRNLERKMEANVDASSEIPGEDAGQPELEKE 1144

## RESULT 12

ITAL\_HUMAN  
ID ITAL\_HUMAN STANDARD; PRT; 1170 AA.  
AC P20701; O43746;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-OCT-2004 (Rel. 45, Last annotation update)  
DE Integrin alpha-L precursor (leukocyte adhesion glycoprotein LFA-1)

alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha chain) (CD11a).  
Name=ITGAL; Synonyms=CD11A;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A. (ISOPFORM 1), AND PARTIAL SEQUENCE.  
RX MEDLINE=89139587; PubMed=2537322;  
RA Larson R.S., Corbi A.L., Berman L., Springer T.;  
RT "Primary structure of the leukocyte function-associated molecule-1 alpha subunit: an integrin with an embedded domain defining a protein superfamily.";  
RL J. Cell Biol. 108:703-712 (1989).  
[2]  
SEQUENCE FROM N.A. (ISOPFORM 2).  
RX MEDLINE=99425270; PubMed=10493829;  
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R., Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L., Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S., Bichler E.B., Harris P.C., Venter J.C., Adams M.D.;  
RT "Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";  
RL Genomics 60:295-308 (1999).  
[3]  
X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.  
RX MEDLINE=96036067; PubMed=7479767;  
RA Qu A., Leahy D.J.;  
RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L beta 2) integrin.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281 (1995).  
[4]  
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.  
RX MEDLINE=96398682; PubMed=8805579;  
RA Qu A., Leahy D.J.;  
RT "The role of the divalent cation in the structure of the I domain from the CD11a/CD18 integrin.";  
RL Structure 4:931-942 (1996).  
[5]  
X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.  
RX MEDLINE=99425288; PubMed=10493852;  
RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G., Cottens S., Weitz-Schmidt G., Hommel U.;  
RT "Structural basis for LFA-1 inhibition upon luvastatin binding to the CD11a I-domain.";  
RL J. Mol. Biol. 292:1-9 (1999).  
CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2, ICAM3 and ICAM4. It is involved in a variety of immune phenomena including leukocyte-endothelial cell interaction, cytotoxic T-cell mediated killing, and antibody dependent killing by granulocytes and monocytes.  
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L associates with beta-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- ALTERNATIVE PRODUCTS:  
CC Event-Alternative splicing; Named isoforms=2;  
CC Name=1;  
CC IsoId=P20701-1; Sequence=Displayed;  
CC Name=2;  
CC IsoId=P20701-2; Sequence=VSP\_002738;  
CC Note=No experimental confirmation available;  
CC TISSUE SPECIFICITY: Leukocytes.  
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins with I-domains do not undergo protease cleavage.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11a entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -



490	RGRVFY	-----ORQLCFEEVSESLQDPGYPLGRFGEAI	TALTDINGDLGVAVGAP	544
529	GEEDREGAVYLPHTSGSIGSPHSORISAGSKLSPLOYFGQSLSGQDLTWGDLVLTIV	588		
545	LEE--QGAYIFNGRHG-GLSPQSPORIGTQVLSGIQWFGRSIHVKDLGGGLADVAV	601		
589	GAGCHVLLRSQVPLRVKAIMENPREVARNVFECDQVV-KGKEAGEVRVCLHVQKSTR	647		
602	GAESQIMVLSRSPVDMVTLMSPSPAEIPVHEVECSYSTSNKMGVNIICFQI-KSLY	660		
648	DLRLEQIGOSVVTYDLALDSGRPHSRVAVNETKNSRRQTQVGLTQTCETLKLQIPNCI	700		
661	PQF-QGELVANLYTILQDCHTRRGLTPGGRHELNRNIAT-TSMCTDFSFFPCV	718		
708	EPVSPIVLRINFSL---VGTPLS--AFGN-----LRPVIABDAQRLFTALPPEKNCGN	757		
719	QDLISPINSLNFSLMEEBGTPRQRAQKOIPIILRPSLHSETWEI-----PPEKNGE	773		
758	DNICQDLSITFSFMSLCLVVGGRFNFVTVVRNDEGDSYRTQVTFPFLDLGYRKVS	817		
774	DKKCEANLEVSPPARSALRUTAFASUSVELSLNLEEDAYVWGLDHFPPGLSFRKVE	833		
818	TLONQORSQWKLACES--ASSTEVSGALKSTCSINHPIPPENSEVFNITFDVDSKAS	875		
834	ML---KPHSQIPVSCBELPEESRLLSRAL---SCNVSPPIKAGHSVALQMMFNTLVNS	887		
876	LGNKLLKANVTSENN---MPTNKTFQLELPVKYAVVMVTVSHGVSTKYNLFTASEN	931		
888	WGDVSVELHANTCNNEEDSLLDENSATTI---IPILYPINILIQDQEDSTLYVSTPKGP	944		
932	TSRVMOHQYOV---SNLQGRSLP-ISLVLFPVRLNQTVIWDROPVTFSENLSSTCHTK-	986		
945	KIHQVKHMYQVRIQPSIHDDHNIPTLEAVVGVPQPPSEGPITHQWSQVMEPPV--PCHVED	1002		
987	ERLPSHSD--FLAELRKAPVNCISIAVCQRIQCDIPFGIOEBFNATLKNLSFDWYIK	1043		
1003	LERLPDAAEPCLPQALFCPPV-----FRQILVQVIGTLELVGEIE	1044		
1044	TSHNHLIVSTAELFNDVSFTLLPGQAFVRSQTEKTVKPEFVENPLPLTVGSSVUGLL	1103		
1045	AS-SMFSLSCSSLSISFNSSKHFLHYGSNASL-AQVVMKDVVYVEKQMLYVLSGIGLL	1102		
1104	LLALITAAVLYKLGFFKQYKQKMMSEG-GPPCAEP	1136		
1103	LLLLIFVLYKVGFKNLKEKMEAGRVNPIP	1136		

[illegible]



QY	417	AMFRQ--NTGMWNSVANVKQTGIGAYFGASLCSVDVDSNGSTDLVLICGAPHYEQTRGO	474
DB	432	LLFQPKRGKGPWSQIEIDGIGSYFGGELCGVDVDRDGETELLIIAALPYGQORGR	491
QY	475	VSVCPPLPGQARWQCDVAVLYCEQGPWGRFCAALTVLGDVNGDKLTDVAIGAPGEEDNR	534
DB	492	VFIY---KQIQLEFQWVSELOQETGYPLGRFGAALTAALTDINGDELTDVAVCAPLEE--Q	546
QY	535	GAIVLPHGTSGSGISPHSHSQRISAGSKLSPRLQYFGQSLSGGQDLTMDGLVDTVTAQGHV	594
DB	547	GAIVYFNGQOG--GLSPRSORIEGTQMFSGIQWFGRSIHGVKDLGGDLADVAVGAEGOV	605
QY	595	LLLRQPVLRVKAIMEFNPVARNVFCNDQVVKKEAG--EVRCLHVQK--STEDRLR	651
DB	606	IVLSSRPVVDIITISVSFPAEIPVHEVECSYSTSNQKKEGVNLTVCQVKSLSI--F	661
QY	652	EQIOGVVYDIALDPSGRPHSAVNETKNSTRROTQVLGLTQTCTBLKLOLPCNIEDPV	711
DB	662	QGHVLANLYTTLQDGHRTSRGLPGGKXKLIGNTAVTPV--KSCFVFWFHPFICIODLI	720
QY	712	SPIVLRINPFL--VGTPLS--AFGNLRPLVLAEDAQRLLFTALFPFKKNCNDNICODDLS	766
DB	721	SPINVSLSYSLWEEECTPRDPRALDRDIPILKPSPHLETKEIPFEKNCOCEDKNCEADLK	780
QY	767	ITPFSMLDCLVVGPPREFNVTVTVNDGEDSYRTQVTFPFDLDSVYKSTLQNSORSOR	826
DB	781	LAFSDMRSKILRLTPSASLSVRLTLRNTAEDAYWVQVTLSPGGLSFRKVEILL--KPHS	837
QY	827	SWRLACESASTEVSGALKSTCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANV	886
DB	838	HVPVGCBEIPPEAVVHS--RALSNCVSPFGEDESMVDIQVWFTLQKSGWDFELQANV	896
QY	887	TS-----ENNMPRTNTEFQLELPKYAVYVTVVSHGVSTKYLNFTASENSTRVNOH	938
DB	897	SCNEDSSLEEDNSATTS-----IPVMYPINVLTKDQENSTLYISFTKSPRIHHVKH	949
QY	939	QYQV---SNLQORSUPLISVLVPLVRLNQTVI---WD---RQVTFSS--ENLSSTCHKE	987
DB	950	IYQVRIQPSNYDNMP--PLEALVRVPRVHSEGLTHKMSIQMEPPVNCSPNLESPSEAE	1008
QY	988	RLPSSHDFLAELKAPVWNCISIAVCQRIQCDIPFGIQEBEFNATLKGNSLFDWYIKTSHN	1047
DB	1009	-----SCSFGT--EPRCFIDF--RQELLVQVNGMVLELRTIKAS--S	1044
QY	1048	HLATVSTAEILFNDVSFTLLPGQAFVRSQETKVEPFEVNPPLIVLGVSGVGLLILAL	1107
DB	1045	MLSCLSSLAISFNSSKHFLHGRNASM--AQVMKVLDLYVEKEMLYLYVLSGIGLLLEFL	1103
QY	1108	ITAALYKLGFPKQYKQDM-----SEGGPPGABPQ	1137
DB	1104	IFIALYKVGFPKRLKEMEANVDASSEIPGEDAGQPELEKE	1145
RESULT 15			
ITAL MOUSE			
ID	ITAL_MOUSE	STANDARD;	PRT; 1163 AA.
AC	P24063;		
DT	01-MAR-1992	(Rel. 21, Created)	
DT	01-MAR-1992	(Rel. 21, Last sequence update)	
DT	05-JUL-2004	(Rel. 44, Last annotation update)	
DE	Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1		
DE	alpha chain) (LFA-1A) (leukocyte function associated molecule 1, alpha		
DE	chain) (CD11a).		
GN	Name=Itgal; Synonyms=Lfa-1;		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91268576; PubMed=2051027;		
RA	Kaufmann Y, Tseng E, Springer T.A.;		
RT	"Cloning of the murine lymphocyte function-associated molecule-1		

Db	897	SCNNEDSSILEDNSATTS-----IPVMYPINVLTKDQENSTLYISFTKSPRIHHVKH	949
QY	939	QYQV-----SNLQORSUPLISVLVPLVRLNQTVI---WD---RQVTFSS--ENLSSTCHTKE	987
Db	950	IYQVRIQPSNYDNMP--PLEALVRVPRVHSEGLTHKMSIQMEPPVNCSPNLESPSEAE	1008
QY	988	RUPSHSDELAELRKAPVWNCISIAVCQRIQCDIPFGIQEBEFNATLKGNSLFDWYIKTSHN	1047
Db	1009	-----SCSFGT--EPRCPIDF--RQELLVQVNGMVLELRTIKAS--S	1044
QY	1048	HLIVSTAEILFNDVSFTLLPGQAFVRSQETKVEPFEVNPDLPLIVGSSVGLLLIAL	1107
Db	1045	MLSCLSSLAISFNSSKHFLHGRNASM--AQVMKVLDLYVEKEMLYLYVLSGIGLLLEFL	1103
QY	1108	ITAALYKLGFPKQYKDM-----SEGGPPGABPQ	1137
Db	1104	IFIALYKVGFPKRLKEMEANVDASSEIPGEDAGQPELEKE	1145
RESULT 14			
AAQ90015			
ID	AAQ90015	PRELIMINARY;	PRT; 1166 AA.
AC	AAQ90015;		
DT	02-MAR-2004	(TrEMBLrel. 27, Created)	
DT	02-MAR-2004	(TrEMBLrel. 27, Last sequence update)	
DT	02-MAR-2004	(TrEMBLrel. 27, Last annotation update)	
DE	Lymphocyte function-associated antigen 1 alpha subunit CD11a.		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovioidea;		
OC	Bovidae; Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;		
RT	"Molecular cloning and sequencing of bovine CD11a.";		
RL	Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AY382558; AAQ90015.1;		
SQ	SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;		
Query Match 26.3%; Score 1546; DB 2; Length 1166;			
Best Local Similarity 34.3%; Pred. No. 9.1e-94;			
Matches 405; Conservative 202; Mismatches 459; Indels 106; Gaps 33;			
QY	1	FNLDTENAMTFOENARG--FGQSVVQLQGSRRVVVGAQPEIIVAAQNRGSLYCDYSTGSC	58
Db	25	YNLDVRHVQNFSPFLAGRHFGYRVLVQ--GNGVVVGAPSE--GNSMGNLYQCQPETGDL	80
QY	59	PIRLQVPEAVNMSIGLSLAATSPQLLACGPTVHQCSENTYVKGICFLFGSLNRQOP	118
Db	81	PVTLTS--SNYTSKYLGMTLATDPTSDNLLACDPGLSRTCDQNTYLSGLCYLIHENLRGPV	138
QY	119	QKFEALRGCPQEDSDIAFLDGGSIIPHDFRMKEMVSTVMEOLKSKSLTFLSLMQYSE	178
Db	139	LQGHFGYQECIKGNVDLVFLFDGMSLQDDEFEKIVDFMKDWMKLSNSYQFAAVQFST	198
QY	179	EFRHFTPEKFNQNNPRSLKIPITQLLGRTHATGLRKVRVLEPNITNGARKNAFKILF	238
Db	199	YFRTEFTLDYIKQKDPDALLAGVKHMLLTNTFGAINVYAVEFRPDLGARPDATKVL	258
QY	239	LLTDEKFGDPLGYEDVPELDRGVIRYVIGVGDAFRSEKRSQELNTVASKPPRDHVQ	298
Db	259	IITDGK---PPTNTTLMRPKTS---RSLLGKGNFKTKESQEAHLHOFASKPVVEEFVKI	311
QY	299	INNEFALKTIQNLREKIFAIGTQTGSSSSPEHEMSQGFSAITSNGLPLSTVGSYDM	358
Db	312	LDTFEKLKDLFTLEKQKIYVIEGTSKQDLTSFNMELSSSGISADLSEGHGVVGAQKDW	371
QY	359	AGGVF--LYTSKEKSTFINNRVDSQNDMDAYLYAA--IILNRVQSLVLGAPYQHIGLV	416
Db	372	AGGFLLDKADKLSSTFGVNGREPLTVESRAGYLVTVTRLPSTRGTMSLLATGAPKYQHVGRV	431

RT alpha-subunit and its expression in COS cells.";  
 RL J. Immunol. 147:369-374 (1991).  
 RN [2].  
 RP SEQUENCE OF 24-42.  
 RX MEDLINE=8518276; PubMed=38871182;  
 RA Springer T.A., Teplov D.B., Dreyer W.J.;  
 RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion  
 RL glycoproteins and unexpected relation to leukocyte interferon.";  
 RL Nature 314:540-542 (1985).  
 CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,  
 CC ICAM3 and ICAM4. Is involved in a variety of immune phenomena  
 CC including leukocyte-endothelial cell interaction, cytotoxic T-cell  
 CC mediated killing, and antibody dependent killing by granulocytes  
 CC and monocytes. Mice expressing a null mutation of the alpha-L  
 CC subunit gene demonstrate impaired tumor rejection and impaired  
 CC leukocytes recruitment.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L  
 CC associates with beta-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Leukocytes.  
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M60778; AAA39426.1; --  
 DR PIR; I56126; I56126.  
 DR HSSP; P20701; LDGO.  
 DR MGD; MGI:96606; Itgal.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 2.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWA; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS0234; VWFA; 1.  
 KW Calcium; Cell adhesion; Direct protein sequencing; Glycoprotein;  
 KW Integrin; Magnesium; Receptor; Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 23  
 FT CHAIN 24 1163  
 FT DOMAIN 24 1087 Integrin alpha-L.  
 FT TRANSMEM 1088 1108 Extracellular (Potential).  
 FT DOMAIN 1109 1163 Potential.  
 FT REPEAT 39 88 Cytoplasmic (Potential).  
 FT REPEAT 89 146 FG-GAP 1.  
 FT DOMAIN 153 325 VWFA.  
 FT REPEAT 348 398 FG-GAP 2.  
 FT REPEAT 399 454 FG-GAP 3.  
 FT REPEAT 455 514 FG-GAP 4.  
 FT REPEAT 516 573 FG-GAP 5.  
 FT REPEAT 576 628 FG-GAP 6.  
 FT REPEAT 628 688 FG-GAP 7.  
 FT CA\_BIND 466 474 Potential.  
 FT CA\_BIND 528 536 Potential.  
 FT CA\_BIND 588 596 Potential.  
 FT SITE 1111 1115 GFFKR motif.  
 FT DISULFID 70 77 By similarity.  
 FT DISULFID 108 126 By similarity.  
 FT DISULFID 147 199 By similarity.  
 FT DISULFID 651 705 By similarity.  
 FT DISULFID 767 773 By similarity.

FT	DISULFID	840	856	By similarity.
FT	DISULFID	993	1009	By similarity.
FT	DISULFID	1017	1048	By similarity.
FT	CARBOHYD	86	86	N-linked (GLCNAC. . .)
FT	CARBOHYD	185	185	N-linked (GLCNAC. . .)
FT	CARBOHYD	270	270	N-linked (GLCNAC. . .)
FT	CARBOHYD	444	444	N-linked (GLCNAC. . .)
FT	CARBOHYD	668	668	N-linked (GLCNAC. . .)
FT	CARBOHYD	696	696	N-linked (GLCNAC. . .)
FT	CARBOHYD	724	724	N-linked (GLCNAC. . .)
FT	CARBOHYD	728	728	N-linked (GLCNAC. . .)
FT	CARBOHYD	776	776	N-linked (GLCNAC. . .)
FT	CARBOHYD	857	857	N-linked (GLCNAC. . .)
FT	CARBOHYD	880	880	N-linked (GLCNAC. . .)
FT	CARBOHYD	899	899	N-linked (GLCNAC. . .)
FT	CARBOHYD	927	927	N-linked (GLCNAC. . .)
FT	CARBOHYD	1056	1056	N-linked (GLCNAC. . .)
SQ	SEQUENCE	1163	AA; 128343	MW; AVA3078489E8232F CRC64;

Query Match 26.0%; Score 1527.5; DB 1; Length 1163;  
 Best Local Similarity 34.1%; Pred. No 1.5e-92;  
 Matches 398; Conservative 217; Mismatches 463; Indels 89; Gaps 36;

QY	1	FNLDTENAMTFQENA-RGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSC	59
Db	24	YNLDRPTQSFQAQGRHFGYQVLQIEDG-VVVGAPGE---GDNTGLVHCRTSSFEQ	79
QY	60	IRLQVPEAVNMSLGLSLAATTPSPQLLAGCTVHTQTSNTYVVKGLCFGLFNSLRQ	119
Db	80	VSLH-GSNHTSKYLGMLTATDAAGSLACDPLGSLTCDQNTYLSGLCLYFQSL	138
QY	120	KPEALRGCPQSDSDIAFLIDGSGSIIIPHDRRMKEWSTVMEQLKSKTLESLMOY	179
Db	139	QNRPAVQECMKGVDLVFLFDGSQLDRKDFKLEFMKDVNRKLSNTSYQFAAVQ	198
QY	180	FRIHFTKFEF-QNNPNRSLIKPIQLLGRTHATGLRKRVRLEFNIENGARKNAFK	238
Db	199	CRTEFTLDYVVKQKNPDLVGLSVQPMELLTNTFPAINVVAHVFKESGARPDATK	258
QY	239	LLTQGEKF--GDPLGYEDVIELDRGVIRVVGVDAPRSEKSRQELNTVASKPRDH	296
Db	259	IITDGEASDKGNI SAHD-----ITRVIIGIGKHFVSQVKTKLHI FASEPVE	309
QY	297	FQINNFEALKTIQNLREKIFAIEGTQGGSSSSEHEMSQEGFSAITNSPLLSVGS	356
Db	310	KILDTFEKLDLFTDLQRIYVAIEGTWRQDITSFNMLSSSGISADUSKHAVVGA	369
QY	357	DWAGGVF-LYTSKEKSTFINNTRVDSMDNAYLGYAAA-IILNRVQSLVGLGAPRY	414
Db	370	DWAGGFLDLREDLOGATFVGQEPILTSVDRGGLGYTVAMWTSRSRPLAAGAPRY	429
QY	415	LVAMFR--QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGTDLVLIGAPHY	472
Db	430	QVLLFQAPAGGRWNQTKIEGTQIGSYFGELCSVDLDDGAEALLIGAPLFFGEQ	489
QY	473	GOVSVCLPRGQARMQCDVLYGEOGPWRFGAALTVLGDVNGDKLTDVAIGAPHE	532
Db	490	GRVFTY---QRRQSLFEMVSELQDGYPLGRFGAALTLDINGDLTDDVAVGAP	545
QY	533	NRGAVYLPHGTSGSGISPSHSQIRAGSKLSPLQYFGQSLSGGQDLMDGLVDLTV	592
Db	546	-QGVAVIENGKPG-GLSPQPSQRIQGAQVFGIRWFGSRHGVKDLGDLADLVV	603
QY	593	HVLLLRSPVLRYKATMEFNPREVARNVFCNDQVVKGEAG-EVVRVCLHVQKSTR	651
Db	604	RVVLSRPRVVDVVTLSFSPEIIPVHEVECSYSAREEQKHGKVLKACRPIKLT	661
QY	652	EQIQSVVTVYDLALDSGRHSRAVFNETKNSITERQTQVLGLTCTCTKLQLEN	711
Db	662	QGRLLANLSTQLDGHMRSGRLFPDGSHELSGNTSITP-DKSCLDLFFHFICIQ	720
QY	712	SPIVLRNFSLV---GTPLSAFGN-LRPVLAEDAQRLEFALPFPEKNCNGNDICQ	767

us-09-902-481b-4.rup

Tue Nov 9 12:56:08 2004

Db 721 SPINVSINFLSEEGTPROKGRAMQPIILRPSIHV-TKEIPKNCGEDKKCEANLTL 779  
QY 768 TFSFMSLDCLVVGGP-----REFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQN 821  
Db 780 SSPARS-----GFLRLMSSASLAVETLSNSGEDAYWRLDLDPRGLSFRKVMQLQ- 831  
QY 822 QRSQRSWRLACESASSTEVSAL-KSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKL 880  
Db 832 --PHSRMPVSCBEL--TEGSSLLTKLKNVSSPIFKAGQEVSLQVMFNTLNSSWEDFV 887  
QY 881 LLKANVTSEN-NMPRTNKTFOLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHQ 939  
Db 888 ELNGTVHCENENSSLQEDNSAATHIPVLPVNLTKBOENSTLYISFTPKGPKTOOVQHV 947  
QY 940 YQVSNLQORSLPISLVFLVPRLNQTIVWDRPQ-----VTFSENLS--TCHTKE-RLP 990  
Db 948 YQV-----RIQFSAYDHNMT-LEALVGVPRHSEDLITYTWSVQTDPLVTCHSEDLKRP 1001  
QY 991 SHSDFLAELRKAPVNVNCSIAVCORIOCDIPFGIOBEFNATLKNLSFDWYIKTSHNHL 1050  
Db 1002 SSE---AEQPCLPGV-----QFRCPVIF---RWEILIQVTGTVELSKEIKAS-STLS 1046  
QY 1051 IVSTABILFNDVSFTLLPGQAFVRSQTEKVEPPEVNPPLPLIVGSSVGGILLALITA 1110  
Db 1047 LCSSLVSFNSKXPHLYGSKA-SEAQVLVKVDLIHEKEMLVHVVLSGIGLVLLFLFL 1105  
QY 1111 ALYKLGFFKQYKQMM-SEGGPPGAEP 1136  
Db 1106 ALYKVGFFRNLRNKEKMEADGGVNGSP 1132

Search completed: November 9, 2004, 12:17:33  
Job time : 142.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds  
(without alignments)  
3950.365 Million cell updates/sec

Title: US-09-902-481B-4  
Perfect score: 5884  
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04.\*

- 1: Geneseqp1980s.\*
- 2: Geneseqp1990s.\*
- 3: Geneseqp2000s.\*
- 4: Geneseqp2001s.\*
- 5: Geneseqp2002s.\*
- 6: Geneseqp2003as.\*
- 7: Geneseqp2003bs.\*
- 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	5852	99.5	1153	2 AAW65090	AAW65090 Human Bet
2	5852	99.5	1153	3 AAB07360	AAB07360 Human CDI
3	5852	99.5	1153	5 AAU80252	AAU80252 Human int
4	5852	99.5	1153	5 ABG61469	ABG61469 Human Bet
5	5852	99.5	1153	5 AA014428	AA014428 Integrin
6	5852	99.5	1153	7 ADD25615	ADD25615 Binding d
7	5842	99.3	1153	2 AAR04136	AAR04136 Alpha sub
8	5836.5	99.2	1152	8 ADM959589	ADM959589 Human int
9	5829.5	99.1	1152	8 ADP12435	ADP12435 Protein e
10	3473	59.0	1163	8 ADP44061	ADP44061 Human CDI
11	3459	58.8	1163	8 ADN02004	ADN02004 Human inf
12	3459	58.8	1163	8 ADQ17510	ADQ17510 Human sof
13	3450	58.6	1163	8 AAR07120	AAR07120 p150.95 a
14	3436	58.4	1163	2 AAW65091	AAW65091 Human Bet
15	3436	58.4	1163	3 AAB07361	AAB07361 Human CDI
16	3436	58.4	1163	5 ABG61470	ABG61470 Human Bet
17	3434	58.4	1163	6 ABU07406	ABU07406 Protein d
18	3434	58.4	1163	7 ADG32005	ADG32005 Human hom
19	3401	57.8	1161	2 AAR78166	AAR78166 Human bet
20	3401	57.8	1161	2 AAW23049	AAW23049 Human bet
21	3401	57.8	1161	2 AAW57491	AAW57491 Human bet
22	3401	57.8	1161	2 AAW65089	AAW65089 Human Bet
23	3401	57.8	1161	2 AAW72825	AAW72825 Human alp
24	3401	57.8	1161	2 AAW73342	AAW73342 Human alp
25	3401	57.8	1161	3 AAB07359	AAB07359 Human alp

26	3401	57.8	1161	5 ABG61468	ABG61468 Human Bet
27	3385.5	57.5	1161	2 AAW23064	AAW23064 Human Bet
28	3385.5	57.5	1161	2 AAW65106	AAW65106 Human Bet
29	3385.5	57.5	1161	2 AAW72837	AAW72837 Human alp
30	3385.5	57.5	1161	2 AAW73343	AAW73343 Human alp
31	3385.5	57.5	1161	3 AAB07376	AAB07376 Human alp
32	3385.5	57.5	1161	5 ABG61485	ABG61485 Human Bet
33	3226.5	54.8	1161	2 AAR78169	AAR78169 Rat beta
34	3224.5	54.8	1161	2 AAW23062	AAW23062 Rat beta
35	3224.5	54.8	1161	2 AAW60004	AAW60004 Rat alpha
36	3224.5	54.8	1161	2 AAW72824	AAW72824 Rat alpha
37	3224.5	54.8	1161	3 AAB07374	AAB07374 Rat alpha
38	3224.5	54.8	1161	5 ABG61483	ABG61483 Rat beta2
39	3217.5	54.7	1161	2 AAW65104	AAW65104 Rat beta2
40	3217.5	54.7	1161	2 AAW73345	AAW73345 Rat alpha
41	3212	54.6	1161	2 AAW23061	AAW23061 Mouse bet
42	3212	54.6	1161	2 AAW60003	AAW60003 Mouse alp
43	3212	54.6	1161	2 AAW65103	AAW65103 Mouse bet
44	3212	54.6	1161	2 AAW72836	AAW72836 Mouse alp
45	3212	54.6	1161	2 AAW73347	AAW73347 Mouse alp

## ALIGNMENTS

## RESULT 1

AAW65090  
ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

XX 28-SEP-1998 (first entry)

XX Human Beta-integrin CD11b subunit protein.

DE Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;  
KW rheumatoid arthritis.

XX Homo sapiens.

XX US5728533-A.

XX 17-MAR-1998.

XX 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

XX 05-AUG-1994; 94US-00286889.

XX 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

XX Van Der Vieren M, Gallatin WM;  
WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using immobilised or  
labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX Example 5; Fig 1A-D; 106pp; English.

XX This sequence represents a human beta-integrin CD11b subunit which is  
used to describe a method for identifying compounds that modulate the  
interaction of the beta-integrin alpha-d subunit with a binding partner  
of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
-d binding partner, one of which is immobilised and the other of which is  
labelled, in the presence of a test compound, and determining if the  
compound affects binding between the alpha-d polypeptide and alpha-d  
binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
comprising the cytoplasmic, transmembrane or extracellular domain of  
alpha-d. Compounds that modulate alpha-d binding could be used to treat

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
CC and rheumatoid arthritis  
XX

SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 2; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGQSVVQLOQSRVVGGAPOEIVAAQORGLSYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARFGQSVVQLOQSRVVGGAPOEIVAAQORGLSYQCDYSTGSCPEI 76  
QY 61 RLQVPEAVNMSLGLSLAATTPPQLLACQPTVHQTCTSENTYVKGCLFLGSGNLRQOPQK 120  
DB 77 RLQVPEAVNMSLGLSLAATTPPQLLACQPTVHQTCTSENTYVKGCLFLGSGNLRQOPQK 136  
QY 121 FPEALRCQPOEDSDIAFLIDGSGSIIIPDPRMKEWSTVMEQLKKSKTLFSLMOYSEEF 180  
DB 137 FPEALRCQPOEDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQKKSKTLFSLMOYSEEF 196  
QY 181 RIHFTFEFQNNPRSLIKEITQLLGRTHATGLRKVVRLEFNITNGARKNAFKILFL 240  
DB 197 RIHFTFEFQNNPRSLVKEITQLLGRTHATGIRKVVRELFNITNGARKNAFKILV 256  
QY 241 TDGEKPGDPLGYEDVPELDBREGVIRVVGDAFRSEKSOELNTVASKRPPRHVFQIN 300  
DB 257 TDGEKPGDPLGYEDVPEADREGVIRVVGDAFRSEKSOELNTVASKRPPRHVFQVN 316  
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360  
DB 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFTNMTVDSDMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFTNMTVDSDMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436  
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGGQSVCP 480  
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTGGQSVCP 496  
QY 481 PRGQARWCCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 540  
DB 497 PRGQARWCCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 556  
QY 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGVAQGHVLLRSQ 600  
DB 557 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTGVAQGHVLLRSQ 616  
QY 601 PVLRVKAIMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLRVKAIMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSRPHSRVFNENKSTRQTVLGTQTCETILKLPNCIEDPVPSPVLRNLF 720  
DB 677 YDLALDSRPHSRVFNENKSTRQTVLGTQTCETILKLPNCIEDPVPSPVLRNLF 736  
QY 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPPEKNGNDNICQDDLSTFTSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPPEKNGNDNICQDDLSTFTSFMSLDCLVVG 796  
QY 781 GPRENVTVTNRDGEDSVRTQVTFPPDLDSYRKVSTLONQORSQSWELACESASTEV 840  
DB 797 GPRENVTVTNRDGEDSVRTQVTFPPDLDSYRKVSTLONQORSQSWELACESASTEV 856  
QY 841 SCALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANVTSENNNPRNKTFE 900  
DB 857 SCALKSTSCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANVTSENNNPRNKTFE 916  
QY 901 QLELPVKYAVYVMTVSHGVSTKYNFTASENTSRVVMQHOYQVSNLQORSLPISLFLVPV 960  
DB 917 QLELPVKYAVYVMTVSHGVSTKYNFTASENTSRVVMQHOYQVSNLQORSLPISLFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNCIAVCQRIQCDIP 1020  
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNCIAVCQRIQCDIP 1036  
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSQTET 1080  
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQCAFVRSQTET 1096  
QY 1081 KVPEFPEVNPDLPLIVGSSVGGLLALLITAAALYKLGFFKQYKDMMSGEGPPGAEPO 1137  
DB 1097 KVPEFPEVNPDLPLIVGSSVGGLLALLITAAALYKLGFFKQYKDMMSGEGPPGAEPO 1153

RESULT 2

AAB07360  
ID AAB07360 standard; protein; 1153 AA.  
XX  
AC AAB07360;  
XX  
DT 17-JAN-2001 (first entry)  
XX  
DE Human CD11b protein sequence.  
XX  
KW Human; macrophage infiltration inhibition; alpha\_d integrin;  
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;  
KW inflammation; leukocyte adhesion deficiency; IAD; Type I diabetes;  
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;  
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;  
KW rheumatoid arthritis; central nervous system injury; CD11b.  
XX  
OS Homo sapiens.  
XX  
PN WO200029446-A1.  
XX  
PD 25-MAY-2000.  
XX  
PF 16-NOV-1999; 99WO-US027139.  
XX  
PR 16-NOV-1998; 98US-00193043.  
PR 08-JUL-1999; 99US-00350259.  
XX  
PA (ICOS-) ICOS CORP.  
XX  
PI Gallatin NW, Van Der Vieren M;  
XX  
XX WPI; 2000-387751/33.  
DR  
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit  
PT macrophage infiltration and reduce inflammation at central nervous system  
PT injury sites.  
XX  
PS Example 5; Fig 1; 270pp; English.  
XX  
CC Integrins are a class of membrane-associated molecules that participate  
CC in cellular adhesion. Integrins are made up of an alpha subunit and a  
CC beta subunit. One class of human integrins are restricted to expression  
CC in white blood cells and have a common beta2 subunit: the leukocyte  
CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins  
CC have an important role in immune and inflammatory responses. The present  
CC protein sequence is the human integrin alpha subunit CD11b. This sequence  
CC was used in an alignment to identify a novel beta2 integrin alpha  
CC subunit: alpha d (AAA60014 and AAB07359). The present sequence has  
CC approximately 60% identity to the protein sequence of alpha d. The  
CC Alpha d gene and protein may be useful in therapy for diseases linked to  
CC aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple  
CC sclerosis, asthma, psoriasis, lung inflammation, acute respiratory  
CC distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency  
CC (IAD). In addition, anti-alpha d monoclonal antibodies may be used in the  
CC inhibition of macrophage infiltration at the site of a central nervous  
CC system injury. The monoclonal antibodies can also be used to detect and  
CC diagnose Crohn's disease  
XX

SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 3; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQBIIVAAQNGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQBIIVAAQNGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPVAVNMSLGLSLAATSPOLLACGPTVHOTCSENTYVKGCLFLFGSNLRQOPQK 120  
DB 77 RLQVPVAVNMSLGLSLAATSPOLLACGPTVHOTCSENTYVKGCLFLFGSNLRQOPQK 136  
QY 121 FPALRCPCPEDSDIAFLIDGSGIIPHDFRRMKEWSTVMEQLKSKTLFSLMQYSEEF 180  
DB 137 FPALRCPCPEDSDIAFLIDGSGIIPHDFRRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196  
QY 181 RIHFTFKFQNNPRSLIKPIITQLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240  
DB 197 RIHFTFKFQNNPRSLVLPITQLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256  
QY 241 TDGEKGDPLGYEDVPELDRGVIRVIGVDAFRSEKROELNVTASPPRDHVFQIN 300  
DB 257 TDGEKGDPLGYEDVPEADREGVIRVIGVDAFRSEKROELNVTASPPRDHVFQVN 316  
QY 301 NFEALKTIONLREKIFAIBGTOTGSSSSPEHEMSQEGFSAITSNGLLSTVGSVDWAG 360  
DB 317 NFEALKTIONLREKIFAIBGTOTGSSSSPEHEMSQEGFSAITSNGLLSTVGSVDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPRYQHILGVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPRYQHILGVAMFR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGSLCSVDVDSNGSTDVLIGAPHYHYEQTRGQGVSVCP 480  
DB 437 QNTGMWESNANVKGTOIGAYFGSLCSVDVDSNGSTDVLIGAPHYHYEQTRGQGVSVCP 496  
QY 481 PRQARWQCDVLYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
DB 497 PRQARWQCDVLYGEQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556  
QY 541 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 600  
DB 557 HGTSGSGISPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLTGVAQGHVLLRSQ 616  
QY 601 PVLRVKAIMFNPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
DB 617 PVLRVKAIMFNPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676  
QY 661 YDLALDSGRPHGRAVENETKNSRRTOTVGLTQTCTETKLQLPNCIEDPVPDIVLRNPF 720  
DB 677 YDLALDSGRPHGRAVENETKNSRRTOTVGLTQTCTETKLQLPNCIEDPVPDIVLRNPF 736  
QY 721 SLVGTPLSAGNLRPVLAEQAQLFTALFFPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAGNLRPVLAEQAQLFTALFFPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
QY 781 GREFNVTIVRNDGSDSTRQTQVTFPDLDSYRKVSTIQNQRQSRWKLACESASTEV 840  
DB 797 GREFNVTIVRNDGSDSTRQTQVTFPDLDSYRKVSTIQNQRQSRWKLACESASTEV 856  
QY 841 SGALKSTCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSSNNPRNTKTF 900  
DB 857 SGALKSTCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSSNNPRNTKTF 916  
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVVMQHYQVSNLQGSRLPISLPLV 960  
DB 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTRVVMQHYQVSNLQGSRLPISLPLV 976  
QY 961 RLNQTVINDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
DB 977 RLNQTVINDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036

QY 1021 FFGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTE 1080  
DB 1037 FFGIOEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTE 1096  
QY 1081 KVEPFEVPNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137  
DB 1097 KVEPFEVPNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1153  
RESULT 3  
ID AAU80252 standard; protein; 1153 AA.  
XX AAU80252;  
XX DT 15-JUL-2002 (first entry)  
XX Human integrin 1 alpha-M subunit protein.  
KW Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;  
KW inflammatory disease; autoimmune disorder; Crohn's disease;  
KW human immunodeficiency virus; HIV; myocardial infarction;  
KW Sjorgen's syndrome; rheumatoid arthritis.  
OS Homo sapiens.  
XX Key Location/Qualifiers  
FH Misc-difference 499.500  
FT /note= "Encoded by GGG CAG AGG"  
XX WO200218583-A2.  
XX PD 07-MAR-2002.  
XX PF 31-AUG-2001; 2001WO-US027227.  
XX PR 01-SEP-2000; 2000US-0229700P.  
XX PA (BLOO-) CENT BLOOD RES INC.  
XX PI Springer TA, Shimoaka M, Lu C;  
XX WPI; 2002-382964/41.  
XX N-PSDB; ABX50046.  
XX Modified integrin-I or integrin I-like domain polypeptide useful as an  
XX immunogen to produce antibodies specific to polypeptide, comprises a  
XX disulfide bond such that polypeptide is stabilized in a desired  
XX conformation.  
XX Disclosure; Page 109-112; 112pp; English.  
XX This invention relates to a modified integrin-I or integrin I-like domain  
XX polypeptide comprising at least one disulfide bond so that the domain is  
XX stabilised in a desired conformation. The polypeptide of the invention  
XX may have antiinflammatory or immunosuppressive activities. The  
XX polypeptides of the invention have an open conformation and are useful as  
XX immunogens to produce antibodies that selectively bind to integrin I-  
XX domain; and for identifying a modulator of integrin activity. Or of  
XX interaction of an integrin and a cognate ligand. The polypeptide of the  
XX invention, or antibodies (preferably anti-IFA-1 antibody) is useful for  
XX treating or preventing an integrin mediated disorder which is an  
XX inflammatory or autoimmune disorder in a subject and for inhibiting the  
XX binding of an integrin to a cognate ligand such as Crohn's disease,  
XX nephritis; human immunodeficiency virus (HIV), myocardial infarction,  
XX Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic  
XX composition comprising the peptide of the invention is useful for  
XX treating an integrin mediated disorder in a subject. The polypeptides  
XX and/or active or antigenic fragments are useful as reagents for diagnosis  
XX of integrin-mediated disorders. The present sequence represents the human  
XX integrin-1 alpha-M protein subunit used to generate the mutant  
XX polypeptides of the invention





SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 5; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAANRGSLYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAANRGSLYQCDYSTGSCPEI	76
QY	61	RLOVPVEAVNMSLGLSLAATSPOLLACGTVHTQTCSENTYVKGCLPFLGNSLRQQPQK	120
DB	77	RLOVPVEAVNMSLGLSLAATSPOLLACGTVHTQTCSENTYVKGCLPFLGNSLRQQPQK	136
QY	121	FPEALRGCPQSDSIAFLIDGSGSIIIDHFRMKEWSTVMEQLKSKTFLSLMOYSEEP	180
DB	137	FPEALRGCPQSDSIAFLIDGSGSIIIDHFRMKEFVSTVMEQLKSKTFLSLMOYSEEP	196
QY	181	RIHFTFEFQNNPNRSLIKPITQLGLRTHPTATGLRKVVRELFNTNGARKNAFKILFL	240
DB	197	RIHFTFEFQNNPNRSLIKPITQLGLRTHPTATGLRKVVRELFNTNGARKNAFKILVVI	256
QY	241	TDGKFGDPLGYEDVIBELDEGVIRVIVGVGDAPRSEKSRQELNTVASXPPRDHVFQIN	300
DB	257	TDGKFGDPLGYEDVIBELDEGVIRVIVGVGDAPRSEKSRQELNTIASKPPRDHVFQVN	316
QY	301	NFEALKTIQNLREKIFAIEGTQTGSSSPSEHEMSQSGFSAATITSGPLLSLVGSDYDAG	360
DB	317	NFEALKTIQNLREKIFAIEGTQTGSSSPSEHEMSQSGFSAATITSGPLLSLVGSDYDAG	376
QY	361	GVFLYTSKSKSTFINMTRVDSMDNDAYLVAAAIILNRNVQSLVLAGAPRYQHIGLVAMFR	420
DB	377	GVFLYTSKSKSTFINMTRVDSMDNDAYLVAAAIILNRNVQSLVLAGAPRYQHIGLVAMFR	436
QY	421	QNTGMWESNANVGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP	480
DB	437	QNTGMWESNANVGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP	496
QY	481	PRGQARWQCDVLYGEOGQWGRFGAALTVDGVNDGKLTVAIGAPGEDNRGAVYLF	540
DB	497	PRGQARWQCDVLYGEOGQWGRFGAALTVDGVNDGKLTVAIGAPGEDNRGAVYLF	556
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVTGAQGHVLLRSQ	600
DB	557	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVTGAQGHVLLRSQ	616
QY	601	PVLVRKAIEMFEPREVARNVFECNDQVVKKEAGEVRVCLHVQKSTRDRRLREGQIQSV	660
DB	617	PVLVRKAIEMFEPREVARNVFECNDQVVKKEAGEVRVCLHVQKSTRDRRLREGQIQSV	676
QY	661	YDLALDSGRPHSAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF	720
DB	677	YDLALDSGRPHSAFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF	736
QY	721	SLVGTPLSAFNLRLPVLAEADQRLFTALPPEKNCNDNIQDGLSITFSFMSLDCLVWG	780
DB	737	SLVGTPLSAFNLRLPVLAEADQRLFTALPPEKNCNDNIQDGLSITFSFMSLDCLVWG	796
QY	781	GPREFNVTVVRNDGDSVRTQVTFPPDLDSVRKUSTIQNQRQSRQSWRLACESASSTEV	840
DB	797	GPREFNVTVVRNDGDSVRTQVTFPPDLDSVRKUSTIQNQRQSRQSWRLACESASSTEV	856
QY	841	SGALKSTCSINHIPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF	900
DB	857	SGALKSTCSINHIPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF	916
QY	901	QLELPVKYAVYVMTVSHGVSTKYLNFNTASNTSRVVMQHQYQVSNLQORSILPISLFLVPV	960
DB	917	QLELPVKYAVYVMTVSHGVSTKYLNFNTASNTSRVVMQHQYQVSNLQORSILPISLFLVPV	976
QY	961	RLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQCDIP	1020
DB	977	RLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCQRIQCDIP	1036

QY	1021	FFGIQEBFNATLKGNSLDFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQGAFFRSQTE	1080
DB	1037	FFGIQEBFNATLKGNSLDFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQGAFFRSQTE	1096
QY	1081	KVEFEFVFNPLPLIVGSSVGLLALLALITAAALYKLGFFKROYKDMMSGGPGGAE	1137
DB	1097	KVEFEFVFNPLPLIVGSSVGLLALLALITAAALYKLGFFKROYKDMMSGGPGGAE	1153

RESULT 5

ID	AAO14428	standard; protein; 1153 AA.
XX	AAO14428;	
AC	AAO14428;	
XX	03-MAY-2002	(first entry)
DT		
XX		
DE		Integrin Mac-1 alpha subunit.
XX		
KW		Mac-1; integrin alpha subunit; variant integrin inserted domain protein; open conformation; integrin related inflammatory disorder;
KW		integrin related immunological disorder; rheumatoid arthritis; ischaemia; reperfusion; hypovolemic shock; infarction; cerebral shock;
KW		viral infection; cancer; gene therapy; vaccine;
KW		bioactive agent screening.

Unidentified.

WO200204521-A2.

17-JAN-2002.

09-JUL-2001; 2001WO-US021805.

07-JUL-2000; 2000US-0216600P.

(CALY ) CALIFORNIA INST OF TECHNOLOGY.

(BLOO-) CENT BLOOD RES.

Springer T;

WPI; 2002-148167/19.

New integrin I domain protein having alteration in at least 2 noncontiguous regions and exits in an open conformation, useful for treating, preventing or suppressing inflammatory or immunological disorders.

Example 1; Fig 1F; 90pp; English.

The invention comprises structurally biased variant integrin inserted (I) domain proteins, wherein the alterations to the protein occur in at least two noncontiguous regions. Specifically the variant integrin I domain proteins are structurally biased to exist in the open conformation, thereby altering the binding ability of the protein. The invention also comprises nucleic acids encoding the variant integrin I domain proteins. The integrin I domain proteins and nucleic acids are useful for treating, preventing or suppressing integrin related inflammatory and immunological disorders (e.g. rheumatoid arthritis). The variant integrin I domain proteins and nucleic acids can also be used for treating: ischaemia/reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral infection; and cancer. The variant integrin I domain nucleic acids and proteins may be used in gene therapy, as vaccines and to screen for bioactive agents. The present amino acid sequence represents the Mac-1 alpha subunit of integrin

SQ Sequence 1153 AA;

Query Match 99.5%; Score 5852; DB 5; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAHQSGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOQPK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOQPK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEMVSTVMEQKKKTLFSLMQYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIIPHDPRMKEMVSTVMEQKKKTLFSLMQYSEEF 196  
QY 181 RIHTTFKEFQNNPNSRLIKPITQLLGRTHTATGLRKVRELFNITNGARKNAFKILFL 240  
DB 197 RIHTTFKEFQNNPNSRLIKPITQLLGRTHTATGLRKVRELFNITNGARKNAFKILFL 256  
QY 241 TDGKFGDPLGYEDVPELDREGVIRYVIGVDGAFRSEKSKQELNTVASKPRDHVFQIN 300  
DB 257 TDGKFGDPLGYEDVPELDREGVIRYVIGVDGAFRSEKSKQELNTVASKPRDHVFQIN 316  
QY 301 NFEALKTIONQREKIFAEGTQTGSSSPFHEMSQEGFSAITNSGPLLSTVGSYDWAG 360  
DB 317 NFEALKTIONQREKIFAEGTQTGSSSPFHEMSQEGFSAITNSGPLLSTVGSYDWAG 376  
QY 361 GFELYTSKEKSTFNMTRVDSMDNDAYLGVAIIILNRVQSLVLGAPRYOHIGLVAMFR 420  
DB 377 GFELYTSKEKSTFNMTRVDSMDNDAYLGVAIIILNRVQSLVLGAPRYOHIGLVAMFR 436  
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTRGGQVSVCP 480  
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYEQTRGGQVSVCP 496  
QY 481 PRGQARWQCDVLYGEGQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRCNAVILP 540  
DB 497 PRGQARWQCDVLYGEGQPMWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRCNAVILP 556  
QY 541 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQHVLRLRSQ 600  
DB 557 HGTSGSGISPSHSORIASKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQHVLRLRSQ 616  
QY 601 PVLVRKALMEPNPREVAENVEFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLVRKALMEPNPREVAENVEFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDGRPHSRVAFNETKSTRQTQVLGTLQTCETLKLQLPNCIEDPVPSPVILRLNF 720  
DB 677 YDLALDGRPHSRVAFNETKSTRQTQVLGTLQTCETLKLQLPNCIEDPVPSPVILRLNF 736  
QY 721 SLVGTPLSAFGLNLRVLAEDAQRLLFTALPPFEKNGCNDNI CODLSITFSFMSLCLVVG 780  
DB 737 SLVGTPLSAFGLNLRVLAEDAQRLLFTALPPFEKNGCNDNI CODLSITFSFMSLCLVVG 796  
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNSQSRWSLACESASSTEV 840  
DB 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNSQSRWSLACESASSTEV 856  
QY 841 SGALKSTCSINHPITPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRINKTEP 900  
DB 857 SGALKSTCSINHPITPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRINKTEP 916  
QY 901 QLELPVKAVYVWVTSYKVLNFTASENTSRVMOHQYQVSNLQORSLPISLFLVFPV 960  
DB 917 QLELPVKAVYVWVTSYKVLNFTASENTSRVMOHQYQVSNLQORSLPISLFLVFPV 976  
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1036  
QY 1021 FFGIQEENATLGNLSFDWYIKTSHNLLIVSTAILFNDSVFTLLPGQGFVRSQTET 1080  
DB 1037 FFGIQEENATLGNLSFDWYIKTSHNLLIVSTAILFNDSVFTLLPGQGFVRSQTET 1096  
QY 1081 KVPEPFEVNPPLIVGSSVGGILLLALITAALYKLGFFKRYQKDMSEGPPGAEPQ 1137

Db 1097 KVPEPFEVNPPLIVGSSVGGILLLALITAALYKLGFFKRYQKDMSEGPPGAEPQ 1153  
RESULT 6  
ADD25615  
ID ADD25615 standard; protein; 1153 AA.  
XX AC ADD25615;  
XX DT 15-JAN-2004 (first entry)  
XX Binding domain-immunoglobulin fusion protein-associated protein #85.  
XX Binding domain; immunoglobulin; fusion protein; cytostatic;  
XX antiarthritic; immunosuppressive; antidiabetic; antichyroid;  
KW neuroprotective; hinge region; immunoglobulin heavy chain;  
KW CH2 constant region; CH3 constant region; IgG1;  
KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
KW type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
XX Unidentified.  
XX OS US2003118592-A1.  
XX PN 26-JUN-2003.  
XX PD 25-JUL-2002; 2002US-00207655.  
XX PF 17-JAN-2001; 2001US-0367358P.  
XX PR 17-JAN-2002; 2002US-00053530.  
XX PR 03-JUN-2002; 2002US-0385691P.  
XX PA (GENE-) GENE-CRAFT INC.  
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
XX WPI; 2003-801317/75.  
XX New binding domain-immunoglobulin fusion protein, useful for treating a  
XX subject having or suspected of having a malignant condition or a B-cell  
XX disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
XX Disclosure; SEQ ID NO 176; 157pp; English.  
XX The invention relates to a binding domain-immunoglobulin fusion protein  
XX comprising a binding domain polypeptide that is fused to an  
XX immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
XX CH2 constant region polypeptide that is fused to the hinge region  
XX polypeptide, and an immunoglobulin heavy chain CH3 constant region  
XX polypeptide that is fused to the CH2 constant region polypeptide. The  
XX hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
XX hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
XX region polypeptide; derived from (a) having 3 or more cysteine residues;  
XX where the mutated human IgG1 immunoglobulin hinge region polypeptide  
XX contains 2 cysteine residues, where the first cysteine is not mutated; a  
XX mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
XX (a) having 3 or more cysteine residues, where the mutated human IgG1  
XX immunoglobulin hinge region polypeptide contains no more than one  
XX cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
XX polypeptide, derived from (a) having 3 or more cysteine residues; where  
XX the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
XX no cysteine residues. The binding domain-immunoglobulin fusion protein is  
XX capable of at least one immunological activity comprising antibody  
XX binding domain polypeptide is capable of specifically binding to an  
XX dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
XX antigen. Also included are an isolated polynucleotide encoding the  
XX binding domain-immunoglobulin fusion protein, a recombinant expression  
XX construct comprising the polynucleotide (operably linked to a promoter),  
XX a host cell transformed or transfected with a recombinant expression  
XX construct, producing the binding domain-immunoglobulin fusion protein, a



FT Modified-site 994..996  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1022..1024  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1045..1047  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1051..1053  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1076..1078  
 FT /label= putative N-glycosylation site  
 FT Region 1106..1134  
 FT /label= putative\_transmembrane\_region  
 XX  
 FN EP364690-A.  
 XX  
 XX 25-APR-1990.  
 XX  
 XX 17-AUG-1989; 89EP-00115159.  
 XX  
 XX 23-AUG-1988; 88US-00235353.  
 PR 09-MAR-1989; 89US-00321239.  
 XX  
 XX (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 XX Springer TA, Corbi A;  
 XX WPI; 1990-125938/17.  
 DR N-PSDB; AAQ04043.  
 XX  
 XX New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating  
 PT inflammation and viral infections, and in diagnosis.  
 PT  
 XX Disclosure; Page ?; 3pp; English.  
 XX  
 XX Mac-1 alpha subunit is involved in the response to inflammation, i.e.  
 CC recognition of and migration to sites of inflammation. It also attaches  
 CC to cellular substrates as part of this function making it useful in  
 CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene  
 CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25  
 CC -MAR-2003 to correct PA field.)  
 CC  
 CC Revised record issued on 09-SEP-2004 : Correction to feature table key  
 CC and pages  
 CC  
 XX Sequence 1153 AA;

Query Match 99.3%; Score 5842; DB 2; Length 1153;  
 Best Local Similarity 99.1%; Pred. No. 0;  
 Matches 1127; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDSYSGCEPI 60  
 Db 17 FNLDTENAMTQENARGQSVVQLQGSRRVVGAPQEIIVAAANQRGSLYQCDSYSGCEPI 76  
 Qy 61 RLQPVVEAVNMSLGLSLAATSPOLLACGPTVHOTCSENTYVVKGLCFGLGSLNRQOPOK 120  
 Db 77 RLQPVVEAVNMSLGLSLAATSPOLLACGPTVHOTCSENTYVVKGLCFGLGSLNRQOPOK 136  
 Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEVSTVMEQLKSKTFLSLMQYSEEF 180  
 Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEVSTVMEQLKSKTFLSLMQYSEEF 196  
 Qy 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVRELFINITNGARKNAFKILILL 240  
 Db 197 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVRELFINITNGARKNAFKILILL 256  
 Qy 241 TDGKFGDPLGVEDVPEADREGVIRYVIGDADFSEKSRQELNTVASKPRDHVFOIN 300  
 Db 257 TDGKFGDPLGVEDVPEADREGVIRYVIGDADFSEKSRQELNTVASKPRDHVFOIN 316  
 Qy 301 NFEALKTIONQUREKIFAIEGTQTGSSSSFEHMSQEGFSAITNGPILLSTVGSDWAG 360  
 Db 317 NFEALKTIONQUREKIFAIEGTQTGSSSSFEHMSQEGFSAITNGPILLSTVGSDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNRYQSLVLGAPRYQHIGLVAMFR 420  
 Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAILRNRYQSLVLGAPRYQHIGLVAMFR 436  
 Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVSCPL 480  
 Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQSVSCPL 496  
 Qy 481 PRQQRARWQCDVLYGEGQGPWRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLFL 540  
 Db 497 PRQQRARWQCDVLYGEGQGPWRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLFL 556  
 Qy 541 HGTSGSGISPSHSORJAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGVAGQHVLRLRSQ 600  
 Db 557 HGTSGSGISPSHSORJAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTGVAGQHVLRLRSQ 616  
 Qy 601 PVLRVKAIMEFNPREVARNVFECDQVVKGKEAGEVRVCLHVOKSTRDRREGIQSVVT 660  
 Db 617 PVLRVKAIMEFNPREVARNVFECDQVVKGKEAGEVRVCLHVOKSTRDRREGIQSVVT 676  
 Qy 661 YDLALDSGRPHSAVFNETKNSRRTQVTLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
 Db 677 YDLALDSGRPHSAVFNETKNSRRTQVTLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736  
 Qy 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
 Db 737 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796  
 Qy 781 GPREFNVTVTRNDGEDSYRTQVTFPPDLQSVKYSTLQNRQSRWRLACESASSTEV 840  
 Db 797 GPREFNVTVTRNDGEDSYRTQVTFPPDLQSVKYSTLQNRQSRWRLACESASSTEV 856  
 Qy 841 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTEP 900  
 Db 857 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPRTNKTEP 916  
 Qy 901 QLELPVKYAVVMVVTSHGVSTKYNLFTASENTSRVMQHVQVSNLQORSPLISLFLVFPV 960  
 Db 917 QLELPVKYAVVMVVTSHGVSTKYNLFTASENTSRVMQHVQVSNLQORSPLISLFLVFPV 976  
 Qy 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
 Db 977 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036  
 Qy 1021 PFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSQTEP 1080  
 Db 1037 PFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGOGAFVRSQTEP 1096  
 Qy 1081 KVEPFEPVNPPLIVGSSVGLLLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1137  
 Db 1097 KVEPFEPVNPPLIVGSSVGLLLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1153  
 RESULT 8  
 ADM99589  
 ID ADM99589 standard; protein; 1152 AA.  
 XX  
 AC ADM99589;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Human integrin alphaM subunit precursor protein.  
 XX  
 KW integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;  
 KW osteopathic; cytostatic; immunosuppressive; antiinflammatory;  
 KW neuroprotective; anticaking; immunotherapy; inflammatory;  
 KW autoimmune disorder; thrombosis; cancer; osteoporosis;  
 KW sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;  
 KW alphaM.  
 XX  
 OS Homo sapiens.  
 XX

FH Key Location/Qualifiers  
 FT Misc-difference 965 /note= "Encoded by CCC"  
 XX WO2004007530-A2.  
 XX 22-JAN-2004.  
 XX 17-JUL-2003; 2003WO-US022301.  
 XX 17-JUL-2002; 2002US-0396783P.  
 XX 17-JUL-2002; 2002US-0396790P.  
 XX 11-SEP-2002; 2002US-0410135P.  
 XX (BLOO-) CENT BLOOD RES INC.  
 XX  
 XX Springer TA, Takagi J;  
 XX WPI; 2004-122877/12.  
 XX N-PSDB; ADM99588.  
 XX  
 XX Novel modified integrin protein having extracellular domains of integrin  
 XX alpha and beta subunits or integrin alpha1 and beta3 subunit, useful for  
 XX treating integrin mediated disorders.  
 XX  
 XX Disclosure; SEQ ID NO 4; 232pp; English.  
 XX  
 XX The invention relates to a novel isolated or recombinant modified  
 XX integrin protein having extracellular domains of integrin alpha and beta  
 XX subunits where one of the subunits has one or more mutations, an altered  
 XX surface feature or an amino acid substitution or internal deletion,  
 XX extracellular domains of the integrin beta subunit that comprise a  
 XX mutation that alters a non-cysteine residue to cysteine or extracellular  
 XX domains of integrin alpha and beta subunits. The polypeptide of the  
 XX invention demonstrates antiproliferative, thrombolytic, anticoagulant,  
 XX osteoprotective, cytostatic, immunosuppressive, antiinflammatory,  
 XX neuroprotective and anticircling activities and may be useful for  
 XX immunotherapy in order to prevent or treat an integrin-mediated disorder  
 XX such as an inflammatory disorder, an autoimmune disorder, thrombosis,  
 XX cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple  
 XX sclerosis. The current sequence is that of the human integrin alpha  
 XX subunit precursor protein of the invention.  
 XX  
 XX Sequence 1152 AA;  
 XX  
 Query Match 99.2%; Score 5836.5; DB 8; Length 1152;  
 Best Local Similarity 99.1%; Pred. No. 0;  
 Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;  
 QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCEPI 60  
 DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCEPI 76  
 QY 61 RLQVPVEAVNMSLGLSLAATSPQLLAGCPTVHQTCSNTYVYKGLCFLFGSNLRQPOK 120  
 DB 77 RLQVPVEAVNMSLGLSLAATSPQLLAGCPTVHQTCSNTYVYKGLCFLFGSNLRQPOK 136  
 QY 121 FPEALRGCPQEDSDIAFLDGSIIIPHDFRRKKEWVSTVMEQLKSKTLFSLMQVSEEF 180  
 DB 137 FPEALRGCPQEDSDIAFLDGSIIIPHDFRRKKEWVSTVMEQLKSKTLFSLMQVSEEF 196  
 QY 181 RIHFTKEFQNNNPRSLIKPIITQLGRTHATGLRKVVRELNTINGARKNAFKLFL 240  
 DB 197 RIHFTKEFQNNNPRSLIKPIITQLGRTHATGLRKVVRELNTINGARKNAFKLFL 256  
 QY 241 TDGEKFGDPLGVEDVPIPELDREGVIRYVIGVDAPFSEKSRQELNTVASKPRDHVFQIN 300  
 DB 257 TDGEKFGDPLGVEDVPIPELDREGVIRYVIGVDAPFSEKSRQELNTVASKPRDHVFQIN 316  
 QY 301 NFEALKTIONQREKIFALEGTQTGSSSSFEHMSQEGFSAATNSGPLLSTVGSYDMAG 360  
 DB 317 NFEALKTIONQREKIFALEGTQTGSSSSFEHMSQEGFSAATNSGPLLSTVGSYDMAG 376

QY 361 GVFLYTSKEKSTFINMTRVSDMNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420  
 DB 377 GVFLYTSKEKSTFINMTRVSDMNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 436  
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSCPL 480  
 DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSCPL 496  
 QY 481 PRGORAPWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDAIGAAGBEDNRGAYL 540  
 DB 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDAIGAAGBEDNRGAYL 555  
 QY 541 HGTSGSGISPSHSQRTAGSKLSPLQYFGOSLGGQDLTMDGLVDLTIVGAQHVVILLRSQ 600  
 DB 556 HGTSGSGISPSHSQRTAGSKLSPLQYFGOSLGGQDLTMDGLVDLTIVGAQHVVILLRSQ 615  
 QY 601 PVLVRKAIMEFNPREVARNVFECDVWVKGKEAGEVRVCLHVOKSTRDRREGIOQSVVT 660  
 DB 616 PVLVRKAIMEFNPREVARNVFECDVWVKGKEAGEVRVCLHVOKSTRDRREGIOQSVVT 675  
 QY 661 YDLALDSGRPHSRAVNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 720  
 DB 676 YDLALDSGRPHSRAVNETKNSTRQTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 735  
 QY 721 SLVGTPLSAPGNLRPVLAEADAQRLFTALPPEKXKCGNDNICQDDLSITFSFMSLDCIVVG 780  
 DB 736 SLVGTPLSAPGNLRPVLAEADAQRLFTALPPEKXKCGNDNICQDDLSITFSFMSLDCIVVG 795  
 QY 781 GPREFNVTVVRNDGSDSVRTQVTFPFLDLSYRKVSTLQNRQSRQSWRLACBSASSTEV 840  
 DB 796 GPREFNVTVVRNDGSDSVRTQVTFPFLDLSYRKVSTLQNRQSRQSWRLACBSASSTEV 855  
 QY 841 SGALKSTCSINHIPIFPENSEVTFNITFDVDSKASLGNKLLKXANVTSENNMPTNKTEF 900  
 DB 856 SGALKSTCSINHIPIFPENSEVTFNITFDVDSKASLGNKLLKXANVTSENNMPTNKTEF 915  
 QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENSTRVWQHGYOVSNLQORSPLISLVLVPV 960  
 DB 916 QLELPVKYAVYVMTVSHGVSTKYLNFTASENSTRVWQHGYOVSNLQORSPLISLVLVPV 975  
 QY 961 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCSIAVCORIOCDIP 1020  
 DB 976 RLNQTVIMDRPOVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVNCSIAVCORIOCDIP 1035  
 QY 1021 FFGIOBEFNATLKGNSFDWYIKTSHNLLIIVSTABILENDSVFTLLPGQGAFFVRSQTET 1080  
 DB 1036 FFGIOBEFNATLKGNSFDWYIKTSHNLLIIVSTABILENDSVFTLLPGQGAFFVRSQTET 1095  
 QY 1081 KVEPEVNPPLIYGVSSVGGILLALITAAALYKLGFFKQYKDMMSGGPGGAEPQ 1137  
 DB 1096 KVEPEVNPPLIYGVSSVGGILLALITAAALYKLGFFKQYKDMMSGGPGGAEPQ 1152  
 RESULT 9  
 ADP12435  
 ID ADP12435 standard; protein; 1152 AA.  
 XX  
 AC ADP12435;  
 XX  
 XX 12-AUG-2004 (firet entry)  
 XX  
 DE Protein encoded by mRNA of the invention #45.  
 XX  
 KW transplant rejection; immune system; rheumatoid arthritis; lupus;  
 KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.  
 OS Homo sapiens.  
 XX  
 XX WO2004042346-A2.  
 XX  
 PD 21-MAY-2004.  
 XX  
 PF 24-APR-2003; 2003WO-US012946.

XX PR 24-APR-2002; 2002US-00131831.  
PR 20-DEC-2002; 2002US-00325899.  
XX PA (EXPR-) EXPRESSION DIAGNOSTICS INC.  
XX PI Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;  
PI Rosenberg S;  
XX WPI; 2004-400724/37.  
XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
XX pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
XX rejection, in an individual, comprises detecting the expression level of  
XX the genes.  
XX Claim 65; SEQ ID NO 2444; 1762pp; English.  
XX The present invention relates to diagnosing or monitoring transplant  
XX rejection, e.g. cardiac or kidney transplant rejection, in an individual  
XX comprises detecting the expression level of one or more genes. The  
XX methods, system and kits are useful in diagnosing or monitoring  
XX transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic  
XX islet, lung, bone marrow or stem cell transplant rejection,  
XX xenotransplant rejection or mechanical organ replacement rejection, in an  
XX individual. The method is also useful in assessing the immune status of  
XX an individual. The methods are also useful in diagnosing and monitoring  
XX diseases that involve the immune system, e.g. rheumatoid arthritis,  
XX lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
XX viral, bacterial or fungal infection. The present sequence represents a  
XX protein that is encoded by the mRNA of the invention.  
XX SQ Sequence 1152 AA;  
Query Match 99.1%; Score 5829.5; DB 8; Length 1152;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;  
QY 1 FNLDTENAMTQENARGQSVQVQGGSRVVGAPQEIIVANQKSLYQCDSYSGCEPI 60  
DB 17 FNLDTENAMTQENARGQSVQVQGGSRVVGAPQEIIVANQKSLYQCDSYSGCEPI 76  
QY 61 RLQVPVEAVNLSGLSLAATTSPPOLLACGPTVHOTCSNTYVYKGLCFGLNSLRQOPQK 120  
DB 77 RLQVPVEAVNLSGLSLAATTSPPOLLACGPTVHOTCSNTYVYKGLCFGLNSLRQOPQK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGIIPHDFFRMKEVSTVMQKSKXTLFLSMQYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDGSGIIPHDFFRMKEVSTVMQKSKXTLFLSMQYSEEF 196  
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRELFTNITNGARKNAFKILPLL 240  
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVRELFTNITNGARKNAFKILVVI 256  
QY 241 TDGEKFGPLGVEDVPELDREGVIRYVIGVDAPRSKROELNTVASKPDRHVFOIN 300  
DB 257 TDGEKFGPLGVEDVPELDREGVIRYVIGVDAPRSKROELNTVASKPDRHVFOIN 316  
QY 301 NFEALKTIONQLREKIFAIEGTQGTSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360  
DB 317 NFEALKTIONQLREKIFAIEGTQGTSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376  
QY 361 GVFLYTSKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVIGAPRYQHIGVAMFR 420  
DB 377 GVFLYTSKSTFINMTRVDSMDNDAYLGAAAIILNRVQSLVIGAPRYQHIGVAMFR 436  
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDNGSTDLVLI GAPHYEQTGRGQVSCPL 480  
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDNGSTDLVLI GAPHYEQTGRGQVSCPL 496  
QY 481 PRGQBARQCDAVLVYGEQQQWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGQBARQCDAVLVYGEQQQWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555

QY 541 HGTSGGSISSHRSORIASGKLSPLQYFGOSLSGGODLTMDGLVLDLTVGAGQHLLRSQ 600  
DB 556 HGTSGGSISSHRSORIASGKLSPLQYFGOSLSGGODLTMDGLVLDLTVGAGQHLLRSQ 615  
QY 601 PVLVRKAIMEFNPREVARNVFECDNVQVKGKEAGEVVRVCLHVQKSTRDLRREGIOQSVVT 660  
DB 616 PVLVRKAIMEFNPREVARNVFECDNVQVKGKEAGEVVRVCLHVQKSTRDLRREGIOQSVVT 675  
QY 661 YDLALDSGRPHSAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 720  
DB 676 YDLALDSGRPHSAVNETKNSRTRQTVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 735  
QY 721 SLVGTPLSAFNGNRPVLAEDAQRILFTALFPFEXKNCNDNICQDDLSITTFSEMSDCLVVG 780  
DB 736 SLVGTPLSAFNGNRPVLAEDAQRILFTALFPFEXKNCNDNICQDDLSITTFSEMSDCLVVG 795  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
DB 796 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 855  
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900  
DB 856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 915  
QY 901 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHQYQVSNLQSRSLPISLVFLVPV 960  
DB 916 QLELPVKYAVVMVTSHGVTSKYLNFTASNTSRVMQHQYQVSNLQSRSLPISLVFLVPV 975  
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
DB 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1035  
QY 1021 PFGIOEEFNATLKNLSFDWIKTSHNHLITVSTAEILFNDVSFTLLPGOGAFVRSQTET 1080  
DB 1036 PFGIOEEFNATLKNLSFDWIKTSHNHLITVSTAEILFNDVSFTLLPGOGAFVRSQTET 1095  
QY 1081 KVEPFEVNPPLPLTVGSSVGGLLLLALITAALYKLGFFKRYKQKMMSEGGPPGAEPPQ 1137  
DB 1096 KVEPFEVNPPLPLTVGSSVGGLLLLALITAALYKLGFFKRYKQKMMSEGGPPGAEPPQ 1152  
RESULT 10  
ADP44061  
ID ADP44061 standard; protein; 1163 AA.  
XX AC ADP44061;  
XX DT 09-SEP-2004 (first entry)  
XX DE Human CD11C protein SEQ ID NO:14.  
XX KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;  
XX KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.  
XX OS Homo sapiens.  
XX FN WO2004053094-A2.  
XX PD 24-JUN-2004.  
XX PF 08-DEC-2003; 2003WO-US039208.  
XX PR 06-DEC-2002; 2002US-0431522P.  
XX PA (PPDP-) PPD DEV LP.  
XX PI Dunn SJ, Holzmayer TA;  
XX PI WPI; 2004-480928/45.  
XX DR N-PSDB; ADP44060.  
XX PT Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

a human host cell useful for preventing and/or treating HIV infection, by identifying an inhibitor of a cell surface polypeptide such as CXCR-4.  
 Claim 1; SEQ ID NO 14; 133pp; English.

The present invention describes a method (M1) for identifying an inhibitor of HIV entry into a human host cell. (M1) comprises identifying an inhibitor of a cell surface polypeptide selected from CXCR-4 (352 amino acids, SEQ ID NO:8, ADP44055), CCR4 (360 amino acids, SEQ ID NO:10, ADP44057), CCR7 (378 amino acids, SEQ ID NO:12, ADP44059), CD11c (1163 amino acids, SEQ ID NO:14, ADP44061), CD47 (323 amino acids, SEQ ID NO:16, ADP44063), CD68 (354 amino acids, SEQ ID NO:18, ADP44065), CD69 (199 amino acids, SEQ ID NO:20, ADP44067), CD74 (566 amino acids, SEQ ID NO:22, ADP44069), CSF3R (836 amino acids, SEQ ID NO:24, ADP44071), RARA (462 amino acids, SEQ ID NO:26, ADP44073), GABBR1 (578 amino acids, SEQ ID NO:28, ADP44075), P2X1 (150 amino acids, SEQ ID NO:30, ADP44077), CD11c (1299 amino acids, SEQ ID NO:32, ADP44079), GPRK6 (576 amino acids, SEQ ID NO:34, ADP44081), or PTK2B (1009 amino acids, SEQ ID NO:36, ADP44083). Also described: (1) an inhibitor of a cell surface polypeptide as described above in a human host cell preventing HIV entry into the human host cell; (2) a pharmaceutical composition comprising an inhibitor of (1) and a carrier; and (3) conferring resistance to HIV infection in an individual, by administering the pharmaceutical composition of (2). An inhibitor of HIV entry has anti-HIV and virucide activities, and can be used as an HIV uptake inhibitor. (M1) is useful for identifying protective compounds that inhibit entry of HIV into cells, useful for the prevention and/or treatment of HIV infection. The present sequence represents human integrin alpha X (CD11c), which is used in the exemplification of the present invention.

Seq Sequence 1163 AA;

Query Match 59.0%; Score 3473; DB 8; Length 1163;  
 Best Local Similarity 61.0%; Pred. No. 1.4e-281;  
 Matches 689; Conservative 142; Mismatches 292; Indels 6; Gaps 4;

QY 1 FNLDENAMTFQENARGFQSVVQGVVAGPQEIIVANRGSLYQCDYSTGCEPI 60  
 DB 20 FNLDTELTAFRVDSDAGFQSVVQVYNSVVVAGPQKITANQGTGLYQGYSTGCEPI 79  
 QY 61 RLQVPVAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVKGCLFCLFGSNLRQOPQK 120  
 DB 80 GLQVPPVAVNMSLGLSLASTSPQLLACGPTVHCECRNXYLTGLCFLGLPT--QLTOR 137  
 QY 121 FPEALRGCPQEDSDIAPLIDGSGIIPHDFRMKEWSTVMEQLKSKTILFSLMQSBEF 180  
 DB 138 LPVSRQCEPQEQDIFVLIDGSGISSRNFAWNVFRAVISQFQRPSTQFSLMQFSNK 197  
 QY 181 RIHFTFKFQNNPNRSLIKPIITQLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240  
 DB 198 QTHFTPEFRSSNPLSLASVHQLQGFTYTATAIQNVVHRLPHASVGARRDAKILVI 257  
 QY 241 TDGEKEDPLGYEDVPELDREGVIRVIGVDAFSEKSRQELNVTASKPPRDHVFQIN 300  
 DB 258 TDGKKKGDSLDYKDVPMADAAIGIYVIGVLAQFQNRNSWKEINDIASKPSQEHIFKVE 317  
 QY 301 NFPAKTIQNLQREKIFALEGTGTGSSSEPEHMSQEGFSAATNSGPLLSTVGSVDWAG 360  
 DB 318 DFDALXIQNLQREKIFALEGTGTGSSSEPEHMSQEGFSAATNSGPLLSTVGSVDWAG 377  
 QY 361 GVFLYTSKESKSTFINNRVDSMDNDAYLGAAAIILNRNVQSLVGLGAPRYQHILGVAMFR 420  
 DB 378 GAFLYPPNMSPTFINNSQENVDRDSVLYGYSTELALWKGVQSLVGLGAPRYQHITGKAVIT 437  
 QY 421 QNTGMHESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQGVSCPL 480  
 DB 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTGQGVSCPL 497  
 QY 481 PRQARWOCDAVLYGEQGPWGRFGAALTVLGDVNGDKLTVDVIGAPGEEDNRGAVILF 540  
 DB 498 PRGWR--RWCDVLYGEQGPWGRFGAALTVLGDVNGDKLTVDVIGAPGEEDNRGAVILF 556  
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTWDGLVDLTGVAQGHVLLLRQ 600

DB 557 HGVLGPSISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTWDGLVDLTGVAQGHVLLLR 616  
 QY 601 PVLRVKALMEFNPREVARNFECDNQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
 DB 617 PVLWVGSMQIPAEIPRSAFECHREQVVEQTLVQSNICLVIDKSKNLLGSRDQSSVT 676  
 QY 661 YDLALDSGRPHSRVAFVETKNSRTRQTVGLTQCTETLKLQLPNCIEDPVSPIVIRLNF 720  
 DB 677 LDALDLPGLSRPRATFOETKNSRVRVGLKAKCFENFLLPSCVEDSVTPIFLRNF 736  
 QY 721 SLVGTPLSAGFNGLPVLAEQAORLFTALFPPEKNCNDNICODDLISITFFSMDLCLVVG 780  
 DB 737 TLVGKPLLAFLNRLPMLAADAQRYFTASLPPEKNCAGDHICQDNLGHSFSPFGLKSLVG 796  
 QY 781 GPRFNVTVTRNDGEDSYRTQVTFPPFLDLVYKRVSTLQNRQSRQWRWLACESASSTEV 840  
 DB 797 SNLELNAEVMVNDGEDSYGTTITFSPAGLSYRYVAEGQKQGLRSLHLTCDAPVG-- 854  
 QY 841 SGALKSTSCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
 DB 855 SQGTWSTSCRINHLIFRGAQITFLATFDVSPRAVLGDRLLLTANVSENNTPTSKTTF 914  
 QY 901 QLELPVAVYVYVTVSHGVSTKYLNFAS--ENTSRVMQHOYVSNLQORSPLISLVLP 959  
 DB 915 QLELPVAVYVTVVSSHEQFTKYLNFSESEKSHVAMHRYQVNNLQORLDPVSNFVWF 974  
 QY 960 VRLNQTVINDRPQVTFSENLSSTCHTKERLPFSDFLAELRKAPVNVNCSTAVCORICDI 1019  
 DB 975 VELNQEAVMVDVESHVPQNPDLRCSSEKIAPPASDFLAHQKPFVLDCCSIAGCLRFCDV 1034  
 QY 1020 PFGIOEENFATLKNLSFDWYIKTSHNHLIVSTABILFENDSVFTLLPQGGAFVRSQTE 1079  
 DB 1035 PFSVQBELDFTLKNLSFGWVRQILQKVVSVVAEITEDTSVYSQLPQGEAFMRAQT 1094  
 QY 1080 TKVEPEVNPPLVIGSSVGGLLILALITAAALKYKGFQKQKDMSE 1128  
 DB 1095 TVLEKYKVNPTPLVIGSSVGGLLILALITAVLYKGVGFKRQYKEMMEE 1143  
 RESULT 11  
 ADN02004  
 ID ADN02004 standard; protein; 1163 AA.  
 XX AC ADN02004;  
 XX DT 17-JUN-2004 (first entry)  
 XX DE Human inflammatory bowel disease marker - CD11c protein.  
 XX KW marker gene; inflammatory bowel disease; IBD; PegammaR IIIa;  
 KW PegammaR IIib; Mig; NRG-2; hexokinase; HM74; REG III; LPAP; Mip-1(beta);  
 KW L-selectin; EGP6; IDO; IL-8; CD11c; TLR2; ulcerative colitis;  
 XX C-rohn's disease.  
 OS Homo sapiens.  
 XX JP2004065120-A.  
 PN 04-MAR-2004.  
 XX 07-AUG-2002; 2002JP-00229705.  
 XX 07-AUG-2002; 2002JP-00229705.  
 XX (SUMU) SUMITOMO SEIYAKU KK.  
 XX WPI; 2004-209124/20.  
 DR N-PSDB; ADN01990.  
 DR Novel marker gene of inflammatory bowel disease (IBD) comprising base  
 PT sequence of e.g., Fc gamma R IIIa, Fc gamma R IIib, Mig, NRG-2,  
 PT hexokinase 3, HM74, CD11c, TLR2 gene, useful for screening therapeutic



PT	agent for IBD.	
XX	Example; SEQ ID NO 27; 151pp; Japanese.	
PS	The invention comprises marker genes for inflammatory bowel disease (IBD)	
XX	- FcgammaR IIa, FcgammaR IIb, Mig, NRG-2, hexokinase, HM74, REG III, LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.	
CC	The DNA and protein sequences of the invention are useful in the	
CC	detection and treatment of IBD (e.g. ulcerative colitis and Crohn's	
CC	disease). The present amino acid sequence represents an IBD marker	
CC	protein of the invention.	
XX	Sequence 1163 AA;	
QY	Query Match 58.8%; Score 3459; DB 8; Length 1163;	
DB	Best Local Similarity 60.9%; Pred. No. 2.1e-280;	
DB	Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;	
QY	1 FNLTENAMTFQENARGGQSVVOLQGSRRVVVVGAPQEIIVAAQNRGSLYQCDSYSGCEPI 60	
DB	20 FNLTDEELTAFRVDSAGFSDSVVQYANVWVVGAPQKITAANOTGGLYQCGYSTGACEPI 79	
QY	61 RLQVFEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFELGNSLRQDQK 120	
DB	80 GLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQCGRNMYLTGLCFELLGPT--QLTQR 137	
QY	121 FPEALRGCPQEDSIAFLIDSGSII PHDFRMKEWSTVMEQLKSKTLESLMOYSEEF 180	
DB	138 LFSVRQECPRQEQDIFVLIDSGSISRRNFATMNFVRAVISQFQPSQFSLMWFQSNKF 197	
QY	181 RIHFFKFEFONNPRSLKIPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240	
DB	198 QHTFFEFRTSNPLSLASVHQLOGFTYATATONVVRHFLHASYGARRDATKILIVI 257	
QY	241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDAPFSEKSRQELNTVASKPRDRHVQIN 300	
DB	258 TDGKKEGSLDYKQVWPADAAGIIRYAIGVLAFONRNSWKELNDAISKPSQEHIFKVE 317	
QY	301 NFEALKTTQNLQREKIFAIEGTQTSSTSSFEHMQEGFSAITNSGPLLSTVGSYDQAG 360	
DB	318 DFDALKDQNLQREKIFAIEGTQTSSTSSFEHMQEGFSAITNSGPLLSTVGSYDQAG 377	
QY	361 GVELTSEKSPFINNTVDSMDNDAYLGAIAAIIILNRKQSLVLGAPRYQHIGLVAMFR 420	
DB	378 GAFLYPPNMSPTFINNSGVNDRSDYLGYSTELALWKGVQSLVLGAPRYQHTGKAVFT 437	
QY	421 QNTGMWESNANVKGTOIGAYFCASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 480	
DB	438 QVSRQWRMKAETGTOIGSYFCASLCSVDVDTGSDTDLVLIGAPHYEQTRGGQVSVCP 497	
QY	481 PRGQARWQCDVLYGEOQWGRFGAALTVLGDVNGDKLTDTVAICAPGEENRGAVILF 540	
DB	498 PRGWR-RWMCDAVLYGEOQWGRFGAALTVLGDVNGDKLTDTVAICAPGEENRGAVILF 556	
QY	541 HTSGSGISPSHSQRIAGSKLSPRLQYFCQSLSGGQDLTMDGLVDLTGCAQGHVLLLRQ 600	
DB	557 HGVLLGPSLPSHSQRIAGSKLSPRLQYFCQALSGGQDLTQDGLVDLAVGARGQVLLRTR 616	
QY	601 PVLRYKAIMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRRLRGQOSVVT 660	
DB	617 PVLWGVSMQFPAEIPRSAFECEQVSEQLTVQSNICLYIDKRSKNLGRDLQSSVT 676	
QY	661 YDLADSGRPHARVFNENKSTRTOVLGHTQTCTELKQLPNCIEPVPPIVLRINF 720	
DB	677 LDALDPGLRSPRAFTQETKNSLSRVRVLGKAHCENFNLLPSCVDSVPTILRLNF 736	
QY	721 SLVGTPLSAFGLNRLPVLAEQAQLFTALFPFEKNCNDNIQQDLSITFSFMSLQCLVVG 780	
DB	737 TLVGKPLLAFLAFLNRLPMLAALQRYPTASLPFEKNCADHICQDNLGIFSFPGLKSLVG 796	
QY	781 GPREFNVTVRNDGDSVTRQVTFPFDLDSVRKYSTLQNRQSRQSRWLAACESASTEV 840	
DB	797 SNLELNAEVMVNDGDSVYGTITTFSHPAGLSYRYVAEGKQKQQLRSLHLTCDSPAVG-- 854	

QY	841 SGALKSTSCSINHPIPPENSEVITFNITFDVDSKASLGNKLLKLVNSENMPRTNKTEF 900	
DB	855 SQGTWSTSCSINHPIFRGGAQITFLATFDVSPRAVLGDRLLLTANVSSENTPRTSKTTF 914	
QY	901 QLELPVKYAVYVMVTSHTKYLINFTAS-ENTSRVMOHQYQVSNLCORSIPISLVFLVP 959	
DB	915 QLELPVKYAVYVTVSSHEQFTKVLNFSESEKESHVAMHRYQVNNLQORLQVSNINFWVP 974	
QY	960 VRLNQTVIDRPOVTFSENLSSCHTKERLPSHSDFLAEILRKAPVNCSTAVCORIOCDI 1019	
DB	975 VELNQEAVMVDVEVSHQPNFSLRCSSEKIAPPASDFLAHQKNPVLDCSTAGCLRFCDV 1034	
QY	1020 PFFGIOEFNATLKNLSFDWYIKTSHNHLIIIVSTAEILFNDSVFTLLPGQGAFAVRQTE 1079	
DB	1035 PFSVQBELDFTLKNLSFGWVRQILQKYSVWSVABITDTSVYSQLPQGEAFMRAQT 1094	
QY	1080 TKVPEFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYKDMSE 1128	
DB	1095 TVLEKYKVNPTPLIVGSSIGGLLLALITAVLYKVGFFKRYKEMME 1143	
RESULT 12		
ID	ADQ17510 standard; protein; 1163 AA.	
XX	ADQ17510;	
XX	26-AUG-2004 (first entry)	
DT	Human soft tissue sarcoma-upregulated protein - SEQ ID 327.	
XX	soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.	
DE	Homo sapiens.	
XX	WO2004048938-A2.	
XX	10-JUN-2004.	
XX	26-NOV-2003; 2003WO-US038193.	
XX	26-NOV-2002; 2002US-0429739P.	
XX	(PROT-) PROTEIN DESIGN LABS INC.	
XX	Aziz N, Ginsburg WM, Zlotnik A;	
XX	WPI; 2004-441208/41.	
XX	Early detection of soft tissue sarcoma comprises determining expression	
XX	of a gene in a first soft tissue sample and a normal soft tissue sample	
XX	and comparing the gene expression, also useful in treating soft tissue	
XX	sarcoma.	
XX	Example 2; SEQ ID NO 327; 210pp; English.	
XX	The invention relates to a novel method for detecting soft tissue sarcoma	
XX	which comprises obtaining a first soft tissue sample from an individual	
XX	and a normal soft tissue sample from the same or different individual,	
XX	determining the expression of a gene in both samples and comparing the	
XX	expression of the gene in both soft tissue samples, where a higher level	
XX	of protein expression in the first soft tissue sample indicates the	
XX	presence of soft tissue sarcoma. The method of the invention has	
XX	cytostatic applications and may be useful for detecting soft tissue	
XX	sarcoma, possibly via gene therapy or vaccine production. The nucleic	
XX	acid sequences may be useful in diagnostic and screening applications.	
XX	The current sequence is that of a human soft tissue sarcoma-upregulated	
XX	protein of the invention. The current sequence is not shown within the	
XX	specification per se but was submitted in CD format by the inventor.	
XX	Sequence 1163 AA;	

Query Match		58.8%;	Score 3459;	DB 8;	Length 1163;
Best local Similarity		60.9%;	Pred. No. 2.le-280;		
Matches 687;		Conservative 142;	Mismatches 294;	Indels 6;	Gaps 4;
QY	1	FNLDTEAMTFOENARFGGVSVOLOQSRVVVVGAPQBIIVANQSGSYQCDYSTGSCPEI	60		
Db	20	FNLDTEELTAFRVDSAGFSDVQYANSVVVVGAPQKITAANQTGGLYQCGYSTGACEPI	79		
QY	61	RLQVPEAVNMISGLSLAANTSPQLACGPTVHQTSENTYVYKGLCFLFGSNLRQPOK	120		
Db	80	GLQVPPPAVNMISGLSLASTTSPQLACGPTVHBEGRNMYLTGLCLLGLPT--QUTQ	137		
QY	121	FPFALRGCPQEDSDIAFLIDSGSIIIPHDFRRMKEWSTVMEQLKKSKTLFSLMOYSEEF	180		
Db	138	LPVSRQECBPQEQDIVFLIDSGSISRNFEATMNFVRAVISQFORPSTQFSLMOFSNKF	197		
QY	181	RIHFTKFEQNNPNRSLIKPIITQLLQRTHTATGLRKVKVRELFINI TNGARKNAFKILFL	240		
Db	198	QTHFTFEFRRTSNPLSLASVHQLQGTVTATAIGNVHRLFHASYGARRDATKILIVI	257		
QY	241	TDGEKFGDPLGYEDVPELDEGVIRVIVGVGDAERSEKSRQELNIVASKPRDRHVFQIN	300		
Db	258	TDGKKGESLDYKDVIPMAADAGIIRVIAIGVGLAFQNRNWKELNDIASPSQEHIFKVE	317		
QY	301	NFEALKTIONQLREKIFAIEGTOTGSSSFEHEMSQEGFSAAITNSGPLLSTVGSYDWAG	360		
Db	318	DFDALKDIONQLREKIFAIEGTETTSSTSSFELEMAQEGFSAVFTPDGVLGAVGSFTWSG	377		
QY	361	GVFLYTSKEKSTINTRVDSNMNDVYLGYAAAIILNRVQSLVGLGAPRYOHIGLVAMER	420		
Db	378	GAFLYPPNMSPTFINNSQENVDRDSYLGYSYTELALWKGVQSLVGLGAPRYOHTGKAVIT	437		
QY	421	QNTGMWESANVKGTOIGAVFGASLCSDVDNSGSDTLVLIGAPHYYEOTRGQGVSVCP	480		
Db	438	QVSRQWRKRAVETGTQIGSTFGASLCSDVDNDSGSDTLVLIGAPHYYEOTRGQGVSVCP	497		
QY	481	PRQARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF	540		
Db	498	PRGWR-RWMCDAVLYGEGQHPWGRFGAALTVLGDVNGDKLTDVIGAPGEENRGAVLYF	556		
QY	541	HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ	600		
Db	557	HGVLGFSISPSHSQRTAGSKLSPRLQYFGQSLGGQDLTQDGLVLAAGARGQVLLRLTR	616		
QY	601	PVLVRKAIEMENPREVARVNECNDQVVKKEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660		
Db	617	PVLWGVSMQFPAEIPRSAFEQREVSEQTLVQSNICLYIDKRSKNLLGSRDLOSVT	676		
QY	661	YDLALDSGRPHRAVENETKNSRRQTQVLGLTQCTETLKLQLPNCIEDPVSPVILRLNF	720		
Db	677	LDLALDPGLSLPRATFOETKNRSLSRVVLGLKAHCENFNLLLPSCVEDSVPTILRLNF	736		
QY	721	SLVGTPLSAFGLNRPVLAEDAQLRFTALPFPFKNCGNDNICODDLSTITFSFMSLDCLVVG	780		
Db	737	TLVGKPLLAFLNRLRPMALAAQRYFTASLFPFKNCGADHICODNLGISFSFPGCLKLLVG	796		
QY	781	GPREFNVTVTRNDGDSYRTOTRPFPLDLSVKYSTLQNBQSQRSWPLACESASTBV	840		
Db	797	SNLELNAEYVWVNDGEDSGTITTFSHPAGLSYRYVAEGQKQQRSLHLTCDAPVVG--	854		
QY	841	SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANYTSENNPRTNKTEF	900		
Db	855	SQGTWSTSCRINHILIFRGAQITFLATFDVSPKAVGLDRLLLTANVSSENNTPRSTKTF	914		
QY	901	QLELPKYAVYVWVTSHGVSSTKYLNPAS-ENTSRVMQHQYQVSNIGQSRSLPISLVLP	959		
Db	915	QLELPKYAVYVWVTSHGVSSTKYLNPAS-ENTSRVMQHQYQVSNIGQSRSLPISLVLP	974		
QY	960	VRLNQTIVDRPQVTFSENLSSTCHTKERLPHSDFLAELRKAPVWVNCISVACQIQCDI	1019		
Db	975	VELNQAENVWDEVSHQNPNSURCSSEKTIAPPASDFLAHQKPNVLDCSIAGCLFRCDV	1034		
QY	1020	PFQIGEEFNATLKGNLSPDWVIKTSNHLIIVSTAELFNDVSFTLLPFGQGFVRSQTE	1079		

Db	1035	PSFSVQELDFTLKNLSFGWVRQILQKKVSVVVAEITFTDTSVYSQLPQGEAFMRAQTT	1094	:     :
----	------	--	------	---

CC Clone lambda X47 was isolated from a cDNA library constructed from total  
 CC RNA extracted from phorbol myristate acetate stimulated HL-60  
 CC myelomonocytic cells. The library was screened with oligonucleotide  
 CC probes based on tryptic peptide fragments of p150.95. The sequence can be  
 CC attached to appropriate control elements and expressed in prokaryotic and  
 CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral  
 CC infection because it interacts with ICAM-1 and inhibits cell-virus  
 CC attachment. It can also be used as an anti-inflammatory agent. See also  
 CC AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003  
 CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)  
 XX  
 SQ Sequence 1163 AA;  
 Query Match 58.6%; Score 3450; DB 2; Length 1163;  
 Best Local Similarity 60.7%; Pred. No. 1.2e-279;  
 Matches 685; Conservative 142; Mismatches 296; Indels 6; Gaps 4;  
 QY 1 FNLDTENAMTFOENARFGQSVQLOGSRVVGAPQIVAAANQSGSLYQCDYSGSCEPI 60  
 Db 20 FNLDTEELTAFRVDSAGFGSVVQYANSWVVGAPQIKITANQGTGGLYQCGYSTGACEPI 79  
 QY 61 RLQVPVAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQDPQK 120  
 Db 80 GLQVPPAVNMSLGLSLASTTSPQLLACGPTVHCEGRNMYLTGLCLFLLGPT--QLTQR 137  
 QY 121 FPEALRGCPQEDSIAFLIDGSGSIIPHDPRMKWSTVMEQLKSKTLFSLMOYSEEF 180  
 Db 138 LPVSRQECPRQEQDIVFLIDGSGSISSRNFAFMNFRVAVISQFQSPSTQFSLMQFSNKF 197  
 QY 181 RIHETVEFQNNPNSRIKIPITOLGTHRTATGLRKVRELFNITGARKNAKFIKILFL 240  
 Db 198 QHTEFEFRFTSNPLSLASVHQGTFTYATAIQNVVHLFRHASYGARDAYKILVI 257  
 QY 241 TDGEKFGDPLGYDVIPELDREGVIRYVIGVDAPRSEKQELNTVASKPPRDRHVQIN 300  
 Db 258 TDGKKEGSLDYKDVIPMADAAGIIRYVIGVGLAFQNRNSWKLNDIASKPSQEHIFKVE 317  
 QY 301 NFEALKITQNLREKIFAIEGTQGTSSSPHEMSQEGFSAITNSGELLSTVSGYSQWAG 360  
 Db 318 DFDALKIQNLQKKEKIFAIEGTETSSSSFELEMAQEGFSAVFTPDGVLGAVGFSFTWSG 377  
 QY 361 GVELYTSKEKSTFINMTVDSDMDAYLGYAAAIILNRNVQSLVGLAPRYQIHGLVAMFR 420  
 Db 378 GAFLYPPNMSTFFINNSQVMDRDSYLYGYSTELALWKGVQSLVGLAPRYQHTGKAVIFT 437  
 QY 421 QNTGMWESNANVKGTCIGAYFGASLCSVDVDSNGSTDVLVLCAPHYVQTRGGQVSVCP 480  
 Db 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTDLVLCAPHYVQTRGGQVSVCP 497  
 QY 481 PRGQARQCDAVLYGEGQPMGRFGAALTIVLDVNGDKLTDTVAIGAPGSDNRGAVYLF 540  
 Db 498 PRGWR-RWMCDAVLYGEGQPMGRFGAALTIVLDVNGDKLTDTVAIGAPGSDNRGAVYLF 556  
 QY 541 HGTSGSISPSHSQRIAGSKLSPRQYFGQSLSGQDLTMDGLVLTGVAQGHVLLRSQ 600  
 Db 557 HGVLGPSISPSHSQRIAGSKLSSRLQYFGQALSGQDLTQDGLVDLAVGARGQVLLRTR 616  
 QY 601 PVLRYKAJMEFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
 Db 617 PVLWGVSNQFTFPIPRSAFCEQVWSEQTLVQSNICLYIDKRSKNLLASRDLOQSSVT 676  
 QY 661 YDLALDGRPHSRVAVNETKSTRQTVLGTQTCETIKLQLPNCIEDPPYSPVLRNLF 720  
 Db 677 LDALDGRSLSPRATFQETKNSLRGRVILGKAHCENFNLLPSCVEDSVPTITRLNF 736  
 QY 721 SLVGTPLGAFGNLRPVLAEDAKRLFTALPFEKNCNDNI CODDLSITFSFMSLCLVVG 780  
 Db 737 TLVHGFLAFLNLRPMLAALAQRYFTASLPFEKNCADHI CODNLGISFSPGLKSLVVG 796  
 QY 781 GPREFNVTVTRNDGEDSVRTQVTFPFLDLVYRKVSTLQNRQSRWSLACSSSTEV 840  
 Db 797 SNLELNAEVMWNWNGEDSVGTITTFHPAGLSVRYVAEGQKQGLRSLHLTCDSA--PVA 854

QY 841 SGALKKSTSCSINHPIFPENSEVFNITFDVDSKASLGNKLLKANKVNTSENMPTNKTEF 900  
 Db 855 SQGTWSTSCSINHPIFRGGAQITFLATFDVSPKAVLGDRLLLANVSSENTERTSKITE 914  
 QY 901 QLELPVKYAVYVMVTVSHGVSTKYLVNTAS-ENTSRVMOHQYQVSNLQORSLSPLSLVLP 959  
 Db 915 QLELPVKYAVYTVVSSHQFTKYLFNSESEKESHVAMHRYQVNNLQGRDLPVSNPWP 974  
 QY 960 VRLNQTVWDRPQVTFSENLSSNCHTKERLPSSHDFLAELRKAPVNVNCIAVCORICDI 1019  
 Db 975 VELNQAVWMDVEVSHPNPNSLRCSSOKIAPPASDFLAHQKPNVLDCSIAGCLFRCDV 1034  
 QY 1020 PFTGIGBEFNATLKGMLSFWDYIKTSHNHLIIVSTAEILFNDSVFTLLPQOGAFVRQTE 1079  
 Db 1035 PFSVQBELDFTLKGMLSGWVQILQKKVSVVVAEITFDTSVYSQLPQGERFMAQTT 1094  
 QY 1080 TKVEPFEVNPPLIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDWSE 1128  
 Db 1095 TVLEKYKVNPTPLIVGSSIGALLLLALITAVLYKVGFFKRYQKEMME 1143  
 RESULT 14  
 AAM65091  
 ID AAM65091 standard; protein; 1163 AA.  
 AC AAM65091;  
 XX 28-SEP-1998 (first entry)  
 DT Human Beta-integrin CD11c subunit protein.  
 XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
 KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;  
 KW rheumatoid arthritis.  
 OS Homo sapiens.  
 XX US5728533-A.  
 PN 17-MAR-1998.  
 PD 07-JUN-1995; 95US-00485618.  
 PF 23-DEC-1993; 93US-00173497.  
 PR 05-AUG-1994; 94US-00286889.  
 PR 21-DEC-1994; 94US-00362652.  
 XX (ICOS-) ICOS CORP.  
 XX Van Der Vieren M, Gallatin WM;  
 XX WPI; 1998-206565/18.  
 DR Screening assay for modulators of integrin binding - using immobilised or  
 PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.  
 PT Example 5; Fig 1A-D; 106pp; English.  
 PS This sequence represents a human beta-integrin CD11c subunit which is  
 XX used to describe a method for identifying compounds that modulate the  
 CC interaction of the beta-integrin alpha-d subunit with a binding partner  
 CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
 CC d binding partner, one of which is immobilised and the other of which is  
 CC labelled, in the presence of a test compound, and determining if the  
 CC compound affects binding between the alpha-d polypeptide and alpha-d  
 CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
 CC comprising the cytoplasmic, transmembrane or extracellular domain of  
 CC alpha-d. Compounds that modulate alpha-d binding could be used to treat  
 CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
 CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
 CC and rheumatoid arthritis  
 XX

SQ Sequence 1163 AA;

Query Match 58.4%; Score 3436; DB 2; Length 1163;  
Best Local Similarity 60.6%; Pred. No. 1.8e-278;  
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;

```
QY 1 FNLDTENAMTFQENARGFGQSVWQSGSRVVVGAPQEIIVAAANQSGLSYQCDYSTGSCPEPI 60
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 20 FNLDTTEELTA PRVDSAGFSDVQVYANSVVVGAPQKIIAANQIGLSYQCGSYTGACEPI 79
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 RLQVPVAVNMGLSLAATSPDLLACGTVHTQTCSENTYVVKGLCFLFGSNLRQPOK 120
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 80 GLQVPPAVNMGLSLASTTSPSOLLACGTVHHEGCRNMYLGLCFLLGPT--QLTQR 137
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 FPEALRCPQDSDIAFLIDSGSIIIPDPRMEKWEVSTWMEQLKSKTFLPSLMQYSEEF 180
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 138 LPVSRQCPQEQDVLIDSGSISRNFAFMNFRAVISQFQRPSTQPSLMQFSNKF 197
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 RIHFTKFEQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 198 QTHFTPEFRRTSNPLSLLASVHQLQGFTYTATAIQNVVHRLFHASGARDALKILIVI 257
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 TDGEKFDPLGYEDVIPDLBREGVIRVIVGVDAFRSEKRQELNVTASKPPRDHVPQIN 300
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 258 TDGKKGEDSLDYKDVIPMDAAGIIRYIAIGVLAQFNENSWKELNDIASKPSQSHIFKVE 317
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 NFPAKTIQNLREKIFAIRGTGTGSSSPSEHEMSQEGFSAITNSGPLLSTVGSYDWAG 360
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 318 DFALKDIONLREKIFAIRGTGTGSSSPSEHEMSQEGFSAITNSGPLLSTVGSYDWAG 377
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILRNVRQSLVILGAPRYQHIGLVAMFR 420
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 378 GAFLYPPNMGPTFINNSQENVMDRDSVILGYSTELALWKGVQSLVGLAPRYQHIGKAVIFI 437
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 QNTGMESNANVGTQIGAFGLASLCSVDVDSNGSTDVLIGAPHYEYQTRGQVSVCP 480
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 438 QVSRQRMKAEBVGTQIGSYFGASLCSVDVDTGSTDVLIGAPHYEYQTRGQVSVCP 497
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 481 PRQARARWQCDVLYGEGQGPWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 498 PRGWR-RWNCDAVLYGEGQGPWRFGAALTVLGDVNGDKLTDVIGAPGEENRGAVYLF 556
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 541 HGTSGSGISPSHSRQIAGSKSLPRLOYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 557 HGVLPGPSISPSHQRQIAGSKSLRQVFGQALSGQDLTODGLVDLAVARGQVILLRTR 616
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 601 PVLRVKALMEFNPVARNVFNCDQVVKGEAGEVVRVCLHVQKSTRDRREGQIQSVVT 660
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 617 PVLWVGVSQFIPAEIIPRSFAECREQVWSEQTILVQSNICLYIDKRSKNLLGSRDLQSSVT 676
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 661 YDLALDSGRPHSAVENETKNSRROTQVGLTQTCTETLQLPNCIEDPVPVILRLNF 720
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 677 LDALAPGLSPRAIIOETKNRSLRVVLGLKAHCENFNLLPSCVEDSVIPILRLNF 736
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 721 SLVGTPLSAPGNLPRVLAEDAQRLLFTALFFPKNCNDNICQDLSITTSFMSLDCLVWG 780
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 737 TLVKGKELLAFNLRPLMALAQRYYFTASUPFERKNCAGHLCQNLGIFSFPGLKSLVVG 796
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 781 GPRFNVTVVRNDGEDSTRTQVTFPPFDLSVRKYVSTIQNQRSQRSWRLACESASTEV 840
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 797 SNLELNAEYVMVNDGSDSYGTTITFSHPAGLSRYRYVAEGKQQLRSLHLC--CSAPVG 854
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 841 SGALKSTCSINHIPEPENSEVFNTITFDVDSKASLGNKLLIKANYTSNNPRTNKTFF 900
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 855 SQGTWSTCSINHLIFRGAQITFLATFDVSPKAVGLDRLLLIANYSSNNIPRTSKTIF 914
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 901 QLELPVKYAVYVWVSHGVSTKYNFTAS-ENTSRVMQHQYQVSNLQGRSLPISLVFLVP 959
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 915 QLELPVKYAVYVWVSHGVSTKYNFTAS-ENTSRVMQHQYQVSNLQGRSLPISLVFLVP 974
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 960 VRLNQTVIWRDQVTFSENLSSTCHTKERLPHSDFLAELRKPAPVNVNCSIAVCQIQCDI 1019
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 975 VELNQAVVMDVEVSHQPNLSRCSSEKIAPPASDPLAHIQKNPVLDCSIAGCLRPRCDV 1034
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
QY 1020 PFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQGFVRSQTE 1079
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1035 PFSVQEELDFTLKNLSFGWVRQILQKQKSVSVSVAEIIFTDSVVSQLPQGEAFMRAQTI 1094
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 1080 TKVEPEFVNPPLPIVGVSSVGGLLLLALITAAALYKLGFFKRYQKDMKSE 1128
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1095 TVLEKYVHPPIPIVGVSSIGGULLLALITAVLYKVGFFKRYQKEMMEE 1143
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

## RESULT 15

```
AAB07361
ID AAB07361 standard; protein; 1163 AA.
XX
AC AAB07361;
XX
DT 17-JAN-2001 (first entry)
XX
DE Human CD11c protein sequence.
XX
KW Human; macrophage infiltration inhibition; alpha d integrin;
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;
KW rheumatoid arthritis; central nervous system injury; CD11c.
XX
OS Homo sapiens.
XX
PN W0200029446-A1.
XX
PD 25-MAY-2000.
XX
PF 16-NOV-1999; 99WO-US027139.
XX
PR 16-NOV-1998; 98US-00193043.
PR 08-JUL-1999; 99US-00350259.
XX
PA (ICOS-) ICOS CORP.
XX
PI Gallatin MW, Van Der Vieren M;
XX
WIPI; 2000-397751/33.
XX
PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
PT macrophage infiltration and reduce inflammation at central nervous system
PT injury sites.
XX
PS Example 5; Fig 1; 270pp; English.
XX
```

Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11c. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has approximately 66% identity to the protein sequence of alpha\_d. The Alpha\_d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease

SQ Sequence 1163 AA;

Query Match 58.4%; Score 3436; DB 3; Length 1163;  
Best Local Similarity 60.6%; Pred. No. 1.8e-278;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
OM protein - protein search, using sw model  
Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds  
(without alignments)  
4119.157 Million cell updates/sec

Title: US-09-902-481B-4  
Perfect score: 5884  
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGAEPPQ 1137

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues  
Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :				Published Applications AA:*	
1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*	2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*	3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*	5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*	6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*	8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*	9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*	11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*	12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*	14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*	15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*	17:	/cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*	18:	/cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*	20:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				* Query Match Length DB ID		Description	
1	5884	100.0	1137	10	US-09-902-481A-4	Sequence 4, Appli	
2	5857	99.5	1137	10	US-09-902-481A-3	Sequence 3, Appli	
3	5852	99.5	1153	9	US-09-350-259-3	Sequence 3, Appli	
4	5852	99.5	1153	10	US-09-902-481A-1	Sequence 1, Appli	
5	5852	99.5	1153	10	US-09-891-943-3	Sequence 3, Appli	
6	5852	99.5	1153	14	US-10-144-259-30	Sequence 30, Appli	
7	5852	99.5	1153	14	US-10-207-655-176	Sequence 176, App	
8	5851	99.4	1137	10	US-09-902-481A-5	Sequence 5, Appli	
9	5845	99.3	1137	10	US-09-902-481A-6	Sequence 6, Appli	
10	5836.5	99.2	1152	9	US-09-945-265-4	Sequence 4, Appli	
11	3459	58.8	1163	14	US-10-116-275-204	Sequence 204, App	
12	3436	58.4	1163	9	US-09-350-259-4	Sequence 4, Appli	
13	3436	58.4	1163	10	US-09-891-943-4	Sequence 4, Appli	

14	3401	57.8	1161	9	US-09-350-259-2	Sequence 2, Appli	
15	3401	57.8	1161	10	US-09-891-943-2	Sequence 2, Appli	
16	3385.5	57.5	1161	9	US-09-350-259-99	Sequence 99, Appl	
17	3385.5	57.5	1161	10	US-09-891-943-99	Sequence 99, Appl	
18	3224.5	54.8	1161	9	US-09-350-259-55	Sequence 55, Appl	
19	3224.5	54.8	1161	10	US-09-891-943-55	Sequence 55, Appl	
20	3212	54.6	1161	9	US-09-350-259-53	Sequence 53, Appl	
21	3212	54.6	1161	10	US-09-891-943-53	Sequence 53, Appl	
22	3203.5	54.4	1151	9	US-09-350-259-37	Sequence 37, Appl	
23	3203.5	54.4	1151	10	US-09-891-943-37	Sequence 37, Appl	
24	3191	54.2	1155	9	US-09-350-259-46	Sequence 46, Appl	
25	3191	54.2	1155	10	US-09-891-943-46	Sequence 46, Appl	
26	1848	31.4	369	13	US-10-087-192-1212	Sequence 1212, Ap	
27	1547.5	26.3	1170	9	US-09-945-265-2	Sequence 2, Appli	
28	1543.5	26.2	1170	15	US-10-261-164-1	Sequence 1, Appli	
29	1520	25.8	1223	16	US-10-408-765A-295	Sequence 295, App	
30	1350.5	23.0	1086	16	US-10-408-765A-1871	Sequence 1871, Ap	
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App	
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App	
33	1141.5	19.4	413	9	US-09-350-259-101	Sequence 101, App	
34	1141.5	19.4	413	10	US-09-891-943-101	Sequence 101, App	
35	1140	19.4	1179	14	US-10-177-550-2	Sequence 2, Appli	
36	1140	19.4	1179	14	US-10-173-551-2	Sequence 2, Appli	
37	1093.5	18.6	1151	10	US-09-884-130-103	Sequence 103, App	
38	1093.5	18.6	1151	10	US-09-836-353A-103	Sequence 103, App	
39	1093.5	18.6	1179	10	US-09-918-715-250	Sequence 250, App	
40	1093.5	18.6	1179	17	US-10-474-794-250	Sequence 250, App	
41	1092	18.6	1188	14	US-10-291-265-810	Sequence 810, App	
42	1088	18.5	1188	14	US-10-291-265-338	Sequence 338, App	
43	1084.5	18.4	1189	10	US-09-884-130-35	Sequence 35, Appl	
44	1084.5	18.4	1189	10	US-09-836-353A-35	Sequence 35, Appl	
45	1084.5	18.4	1189	15	US-10-262-839-4	Sequence 4, Appli	

ALIGNMENTS

RESULT 1  
US-09-902-481A-4  
; Sequence 4, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimooka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902.481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-4

Query Match	100.0%;	Score 5884;	DB 10;	Length 1137;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1137;	Conservative	0;	Mismatches	0;
			Indels	0;
			Gaps	0;
Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVQAPQEIIVAAORGSLYQCDYTGSCPEI	60	
Db	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVQAPQEIIVAAORGSLYQCDYTGSCPEI	60	
Qy	61	RLQVPVAVNMVSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK	120	
Db	61	RLQVPVAVNMVSLGLSLAATTSPPQLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK	120	

```
QY 121 PPEALRGCPQEDSDIAFLIDSGSGIIPHDPRMKWVSTVMEOLKSKXTLFSLMQYSEEF 180
DB 121 PPEALRGCPQEDSDIAFLIDSGSGIIPHDPRMKWVSTVMEOLKSKXTLFSLMQYSEEF 180
QY 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILFL 240
DB 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILFL 240
QY 241 TDGKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSOELNTVASKPPRDHVFQIN 300
DB 241 TDGKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSOELNTVASKPPRDHVFQIN 300
QY 301 NFEALKTQNLREKIFAEGTQTCSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTQNLREKIFAEGTQTCSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQVSVCP 480
DB 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQVSVCP 480
QY 481 PRGQARWQCCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
QY 601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSCRPHSRAVFNETKNSSTRQTQVLGLTQTCETILKLPNCIEDPVSPIVRLNPF 720
DB 661 YDLALDSCRPHSRAVFNETKNSSTRQTQVLGLTQTCETILKLPNCIEDPVSPIVRLNPF 720
QY 721 SLVGTPLSAFQNLRLVLAEDAQRLETFALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFQNLRLVLAEDAQRLETFALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
QY 841 SGALKSTCSINHPIFFPENSEVTENITFDVDSKASIGNKLLKANVTSENMPRTNKTEF 900
DB 841 SGALKSTCSINHPIFFPENSEVTENITFDVDSKASIGNKLLKANVTSENMPRTNKTEF 900
QY 901 QLELPVKYAVVMVTSHGVTSKYLNFTASENTSRVQHQQYQVSNLQSRSLPISLVFLVPV 960
DB 901 QLELPVKYAVVMVTSHGVTSKYLNFTASENTSRVQHQQYQVSNLQSRSLPISLVFLVPV 960
QY 961 RLNQTVIWDROQVTFSENLSSTCHTKERLPSHDSFLAELRKA PVNCSIAVQRIQCDIP 1020
DB 961 RLNQTVIWDROQVTFSENLSSTCHTKERLPSHDSFLAELRKA PVNCSIAVQRIQCDIP 1020
QY 1021 PFGIQQEENATLKGNSLFDWYIKTSHNLLLIYSTAEILFNDVSFTLLPGQGFVRSQTET 1080
DB 1021 PFGIQQEENATLKGNSLFDWYIKTSHNLLLIYSTAEILFNDVSFTLLPGQGFVRSQTET 1080
QY 1081 KVEPEVPNPLPIVSSVGGLLALLALITAALYKLGFFKRYQKMDMSEGGPGABPQ 1137
DB 1081 KVEPEVPNPLPIVSSVGGLLALLALITAALYKLGFFKRYQKMDMSEGGPGABPQ 1137
```

RESULT 2

US-09-902-481A-3

; Sequence 3, Application US/09902481A

```
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-3
```

```
Query Match 99.5%; Score 5857; DB 10; Length 1137;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1131; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVOLQSGRVVVGAPQETVAANORGLYQCDYSTGSCBPI 60
DB 1 FNLDTENAMTFOENARGFGQSVVOLQSGRVVVGAPQETVAANORGLYQCDYSTGSCBPI 60
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGTPVHTQCSNTYVVKGLCFLFGSNLRQOPK 120
DB 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGTPVHTQCSNTYVVKGLCFLFGSNLRQOPK 120
QY 121 FPEALRGCPQEDSDIAFLIDSGSGIIPHDPRMKWVSTVMEOLKSKXTLFSLMQYSEEF 180
DB 121 FPEALRGCPQEDSDIAFLIDSGSGIIPHDPRMKWVSTVMEOLKSKXTLFSLMQYSEEF 180
QY 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILFL 240
DB 181 RIHFTFKFQNNPNSRLIKPITQLLGRTHATGLRKVVRELFINITNGARKNAFKILFL 240
QY 241 TDGKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSOELNTVASKPPRDHVFQIN 300
DB 241 TDGKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSOELNTVASKPPRDHVFQIN 300
QY 301 NFEALKTQNLREKIFAEGTQTCSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTQNLREKIFAEGTQTCSSSFHEMSQEGFSAAITNSGPLLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINMTVDSDMDNDAYLGYAAAIILNRVQSLVLAGAPRYQHIGLVAMFR 420
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQVSVCP 480
DB 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTGGQVSVCP 480
QY 481 PRGQARWQCCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
QY 601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRVKAIEMFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
QY 661 YDLALDSCRPHSRAVFNETKNSSTRQTQVLGLTQTCETILKLPNCIEDPVSPIVRLNPF 720
DB 661 YDLALDSCRPHSRAVFNETKNSSTRQTQVLGLTQTCETILKLPNCIEDPVSPIVRLNPF 720
```



QY 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
Db 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
Db 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
QY 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Db 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
QY 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSISLVFLVPV 960  
Db 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSISLVFLVPV 960  
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020  
Db 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020  
QY 1021 FFGIQQEENATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080  
Db 1021 FFGIQQEENATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080  
QY 1081 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQDMSEGPPGAEPQ 1137  
Db 1081 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQDMSEGPPGAEPQ 1137

RESULT 3  
US-09-350-259-3  
; Sequence 3, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20020062008A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER FILING DATE: 1999-07-08  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-3

Query Match 99.5%; Score 5852; DB 9; Length 1153;  
Best local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVWOLQSRVVVVGAPQEIIVANORGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFOENARFGQSVWOLQSRVVVVGAPQEIIVANORGSLYQCDYSTGSCPEI 76  
QY 61 RLOQPVAVNMSLGLSLAATSPPLQIACGPTVHOTCSENTYVKGICFLFGSNLRQOPQK 120  
Db 77 RLOQPVAVNMSLGLSLAATSPPLQIACGPTVHOTCSENTYVKGICFLFGSNLRQOPQK 136  
QY 121 FPEALRGCPQSDSDIAFLIDSGSIIIPHDFRMRKEFVSTYMEQLKSKTFLSLMOYSEEF 180

Db 137 FPEALRGCPQSDSDIAFLIDSGSIIIPHDFRMRKEFVSTYMEQLKSKTFLSLMOYSEEF 196  
QY 181 RIHFTFEFQNNPNRSLIKPITQLLORTHTATGLRKVRELFINITNGARKNAKFLPLL 240  
Db 197 RIHFTFEFQNNPNRSLIKPITQLLORTHTATGLRKVRELFINITNGARKNAKFLVVI 256  
QY 241 TDGEKFGDPLGYEDVIPELDREGVIRVYVIGVDGAFRSEKSRQELNLTVASKEPRDHVPQIN 300  
Db 257 TDGEKFGDPLGYEDVIPELDREGVIRVYVIGVDGAFRSEKSRQELNLTVASKEPRDHVPQIN 316  
QY 301 NFEALKTIQNLREKIFAIEGTQGTGSSSFEHEMSQEGFSAAITSNGLPILLSTVGSYDWAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTQGTGSSSFEHEMSQEGFSAAITSNGLPILLSTVGSYDWAG 376  
QY 361 GVFLYTSKESKSTFINMTVRVSDMNDAYLGYAAAIILNRVQSLVLGAPRQHIHGLVAMFR 420  
Db 377 GVFLYTSKESKSTFINMTVRVSDMNDAYLGYAAAIILNRVQSLVLGAPRQHIHGLVAMFR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSCPL 480  
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSCPL 496  
QY 481 PRGQARWQCDAYLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVLYF 540  
Db 497 PRGQARWQCDAYLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEEDNRGAVLYF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQHVLRLSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAQHVLRLSQ 616  
QY 601 PVLRVKAIMEFNPREVARNVFECNDQVWKGKEAGEVRVCLHVOKSTRDRRREGOIQSVVT 660  
Db 617 PVLRVKAIMEFNPREVARNVFECNDQVWKGKEAGEVRVCLHVOKSTRDRRREGOIQSVVT 676  
QY 661 YDLALDSGRPHSRAVFNETKNSTRRTQVGLTQTCETLKLQCPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRAVFNETKNSTRRTQVGLTQTCETLKLQCPNCIEDPVSPIVLRNF 736  
QY 721 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIPPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916  
QY 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSISLVFLVPV 960  
Db 917 QLELPVKYAVYVMVTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSLSISLVFLVPV 976  
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1036  
QY 1021 FFGIQQEENATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080  
Db 1037 FFGIQQEENATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQAFVRSQTET 1096  
QY 1081 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQDMSEGPPGAEPQ 1137  
Db 1097 KVEPPEVNPPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQDMSEGPPGAEPQ 1153

RESULT 4  
US-09-902-481A-1  
; Sequence 1, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy

```
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; CURRENT FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1153
; TYPE: PR
; ORGANISM: Homo sapiens
; NAME/KEY: mat_peptide
; LOCATION: (17)..()
; OTHER INFORMATION:
US-09-902-481A-1

Query Match 99.5%; Score 5852; DB 10; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFGLFSGNLROQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFGLFSGNLROQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKKSKTLFSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256

QY 241 TDGEXFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGEXFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQIN 316

QY 301 NPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINNRVDSMDNDAYLGAAAIILRNQVSLVGLAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNRVDSMDNDAYLGAAAIILRNQVSLVGLAPRYOHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTIQAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQSVCLP 480
Db 437 QNTGMWESNANVKGTIQAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQSVCLP 496

QY 481 PRGQARWQCDVLYGEOGQPWGRFGAALTVLGDVNGDKLTVAIGAPGEENRGAIVLF 540
Db 497 PRGQARWQCDVLYGEOGQPWGRFGAALTVLGDVNGDKLTVAIGAPGEENRGAIVLF 556

QY 541 HGTSGSGISPSHSORIASKLSPLRQYFGQSLSGGDLTMDGLVLDLTVGAQGHVLLLRSQ 600
Db 557 HGTSGSGISPSHSORIASKLSPLRQYFGQSLSGGDLTMDGLVLDLTVGAQGHVLLLRSQ 616

QY 601 PVLRVKAIIMEFNPVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIIMEFNPVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHRAVENETKSTRQTQVLGTLTQTCETLKLQIPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHRAVENETKSTRQTQVLGTLTQTCETLKLQIPNCIEDPVSPVILRLNF 736

721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKCKGNDNICODDLSTITFSWMLDCLVVG 780
737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKCKGNDNICODDLSTITFSWMLDCLVVG 796
781 GPREFNVTVVRNDGEDSYRTQVTFFFPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
797 GPREFNVTVVRNDGEDSYRTQVTFFFPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMMPRNKTEF 900
857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMMPRNKTEF 916
901 QLELPVKYAVVMVTSHGVTSTKYNFTASENTSRVMQHOYQVSNLQSRSLPISLVFLVPV 960
917 QLELPVKYAVVMVTSHGVTSTKYNFTASENTSRVMQHOYQVSNLQSRSLPISLVFLVPV 976
961 RLNOTVIWDRPOVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020
977 RLNOTVIWDRPOVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036
1021 PFGIOEBENATLKGNSLSEWDYIKTSHNHLLIVSTAEILLPNDVSFTLLPGQGAFFVSQDET 1080
1037 PFGIOEBENATLKGNSLSEWDYIKTSHNHLLIVSTAEILLPNDVSFTLLPGQGAFFVSQDET 1096
1081 KVEPEFVNPPLIVGSSVGGLLLLALITALYKLGFFKROYKDMMSGGPPGAPQ 1137
1097 KVEPEFVNPPLIVGSSVGGLLLLALITALYKLGFFKROYKDMMSGGPPGAPQ 1153

RESULT 5
US-09-891-943-3
; Sequence 3, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1 Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PR
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.5%; Score 5852; DB 10; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76

QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFGLFSGNLROQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHTCSENTYVVKGLCFGLFSGNLROQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKKSKTLFSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHATGLRKVVRELFNITNGARKNAFKILVVI 256

QY 241 TDGEXFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQIN 300
Db 257 TDGEXFGDPLGVEDVPEADREGVIRYVIGVDAPRSEKSRQELNTVASKPPRDHVFQIN 316

QY 301 NPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINNRVDSMDNDAYLGAAAIILRNQVSLVGLAPRYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINNRVDSMDNDAYLGAAAIILRNQVSLVGLAPRYOHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTIQAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQSVCLP 480
Db 437 QNTGMWESNANVKGTIQAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQSVCLP 496

QY 481 PRGQARWQCDVLYGEOGQPWGRFGAALTVLGDVNGDKLTVAIGAPGEENRGAIVLF 540
Db 497 PRGQARWQCDVLYGEOGQPWGRFGAALTVLGDVNGDKLTVAIGAPGEENRGAIVLF 556

QY 541 HGTSGSGISPSHSORIASKLSPLRQYFGQSLSGGDLTMDGLVLDLTVGAQGHVLLLRSQ 600
Db 557 HGTSGSGISPSHSORIASKLSPLRQYFGQSLSGGDLTMDGLVLDLTVGAQGHVLLLRSQ 616

QY 601 PVLRVKAIIMEFNPVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIIMEFNPVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHRAVENETKSTRQTQVLGTLTQTCETLKLQIPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHRAVENETKSTRQTQVLGTLTQTCETLKLQIPNCIEDPVSPVILRLNF 736
```

```
197 RIHFTKEFQNNPNSRLVKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
241 TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGAPRSEKSKQELNIVASKPRDRHVFOIN 300
257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAPRSEKSKQELNITASKPRDRHVFOVN 316
301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 480
437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 496
481 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 540
497 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 556
541 HGTSGSGISPSHQSRIAGSKLSPRLQVFGQSLSGGQDLTMDGLVDLTGAGHVLRLRSQ 600
557 HGTSGSGISPSHQSRIAGSKLSPRLQVFGQSLSGGQDLTMDGLVDLTGAGHVLRLRSQ 616
601 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 660
617 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 676
661 YDLALDSGRPHSRAVFNETKNSTRQTQVGLGTOTCETLKLQPLNCTEDPVSPIVLRINF 720
677 YDLALDSGRPHSRAVFNETKNSTRQTQVGLGTOTCETLKLQPLNCTEDPVSPIVLRINF 736
721 SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
737 SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
781 GPREFNVTVVRNDGDSYRTQVTFPPDLISYKRVSTLQNRQSRWRLACESASSTEV 840
797 GPREFNVTVVRNDGDSYRTQVTFPPDLISYKRVSTLQNRQSRWRLACESASSTEV 856
841 SGALKSTSCSINHPIFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
857 SGALKSTSCSINHPIFENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
901 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHQVQSNLQORSPLISLVFLVPV 960
917 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMQHQVQSNLQORSPLISLVFLVPV 976
961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVGSIIVCQRIQCDIP 1020
977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNVGSIIVCQRIQCDIP 1036
1021 FFGIOEBFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSTPLPGOGAFVRSOTET 1080
1037 FFGIOEBFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSTPLPGOGAFVRSOTET 1096
1081 KVPEPFVNPPLIVGSSVGLLALLALITAAALYKLGFFKRYQKDMMSGEGPPGAEPQ 1137
1097 KVPEPFVNPPLIVGSSVGLLALLALITAAALYKLGFFKRYQKDMMSGEGPPGAEPQ 1153
```

## RESULT 6

```
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
```

```
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144, 259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758, 493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221, 950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30
```

```
Query Match 99.5%; Score 5852; DB 14; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 1 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 76
Qy 61 RLQVVEAVNMVSLGLSLAATTSPPOLACGPTVHOTCSENTYVKGCLFGLFGSNLRQPOK 120
Db 77 RLQVVEAVNMVSLGLSLAATTSPPOLACGPTVHOTCSENTYVKGCLFGLFGSNLRQPOK 136
Qy 121 FPEALRCCPQEDSDIAFLIDGSGSIIIPHDFRRMKEMVSTVMEQLKKSKTLFSLMOYSEEF 180
Db 137 FPEALRCCPQEDSDIAFLIDGSGSIIIPHDFRRMKEFVSTVMEQLKKSKTLFSLMOYSEEF 196
Qy 181 RIHFTKEFQNNPNSRLVKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 240
Db 197 RIHFTKEFQNNPNSRLVKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
Qy 241 TDEKFGDPLGYEDVPELDREGVIRYVIGVDGAPRSEKSKQELNIVASKPRDRHVFOIN 300
Db 257 TDEKFGDPLGYEDVPEADREGVIRYVIGVDGAPRSEKSKQELNITASKPRDRHVFOVN 316
Qy 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 496
Qy 481 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 540
Db 497 PRGORARWQCDVAVLYGEGQGPWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNREGAVYLF 556
Qy 541 HGTSGSGISPSHQSRIAGSKLSPRLQVFGQSLSGGQDLTMDGLVDLTGAGHVLRLRSQ 600
Db 557 HGTSGSGISPSHQSRIAGSKLSPRLQVFGQSLSGGQDLTMDGLVDLTGAGHVLRLRSQ 616
Qy 601 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 660
Db 617 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDRLEGOIQSVVT 676
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVGLGTOTCETLKLQPLNCTEDPVSPIVLRINF 720
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVGLGTOTCETLKLQPLNCTEDPVSPIVLRINF 736
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796
Qy 781 GPREFNVTVVRNDGDSYRTQVTFPPDLISYKRVSTLQNRQSRWRLACESASSTEV 840
```

Db 797 GPREFNVTVVRNDGEDSYRTQVTFFFPDLDSYRKVSTLQNRQSRWLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916  
Qy 901 QLELPVKYAVVMVTSYKYNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 960  
Db 917 QLELPVKYAVVMVTSYKYNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 976  
Qy 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036  
Qy 1021 PFGIOEEFNATLKGNSLSDWYKTSYKYNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 1080  
Db 1037 PFGIOEEFNATLKGNSLSDWYKTSYKYNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 1096  
Qy 1081 KVEPEVNPPLIIVGSSVGGLLLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1137  
Db 1097 KVEPEVNPPLIIVGSSVGGLLLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1153

RESULT 7  
US-10-207-655-176  
; Sequence 176, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 39069, 401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655  
; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 176  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-207-655-176

Query Match 99.5%; Score 5852; DB 14; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFQGSVVQLQGRVVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFOENARGFQGSVVQLQGRVVVVGAPQEIIVAAQNRGLYQCDYSTGSCPEI 76

Qy 61 RLQVPEAVNNSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQOPK 120  
Db 77 RLQVPEAVNNSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQOPK 136

Qy 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEWVSTVMEQLKKSKTLFSLMOYSEEF 180  
Db 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKKEWVSTVMEQLKKSKTLFSLMOYSEEF 196

Qy 181 RIHFTFKFQNNPNRSLIKPTQLLGRTHATAGLRKVRELFNITNGARKNAFKLLFLL 240  
Db 197 RIHFTFKFQNNPNRSLIKPTQLLGRTHATAGLRKVRELFNITNGARKNAFKLLFLL 256

Qy 241 TDEKFGDPLGYEDVPELDRGVTRYVIGVDGAFRSEKSRQBLNTVASKPRDHFQIN 300  
Db 257 TDEKFGDPLGYEDVPELDRGVTRYVIGVDGAFRSEKSRQBLNTVASKPRDHFQIN 316

Qy 301 NFEALXTQKREKIFALEGTQSGSSSFEHMSQEGFSAAITSGNPLSTVGSYDMWAG 360  
Db 317 NFEALXTQKREKIFALEGTQSGSSSFEHMSQEGFSAAITSGNPLSTVGSYDMWAG 376

Qy 361 GVFLYTSKESKTFINMTRVDSMDNDAYLGAAAILLRNVQSLVLCAPRYOHIGLVAMER 420  
Db 377 GVFLYTSKESKTFINMTRVDSMDNDAYLGAAAILLRNVQSLVLCAPRYOHIGLVAMER 436

## RESULT 8

US-09-902-481A-5  
; Sequence 5, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic

Qy 421 QNTGMWBSNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 480  
Db 437 QNTGMWBSNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSVCP 496  
Qy 481 PRQGRARWQCDVAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 540  
Db 497 PRQGRARWQCDVAVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 556  
Qy 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDQLTMDGLVDLTIVGAQGHVLLLSRQ 600  
Db 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLSGQDQLTMDGLVDLTIVGAQGHVLLLSRQ 616  
Qy 601 PVLVKALIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 617 PVLVKALIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRVAFNETKNSSTRQTVGLTQTCETLKLQLPNCIEDPSPVLRNLNF 720  
Db 677 YDLALDSGRPHSRVAFNETKNSSTRQTVGLTQTCETLKLQLPNCIEDPSPVLRNLNF 736  
Qy 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFFFPDLDSYRKVSTLQNRQSRWLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFFFPDLDSYRKVSTLQNRQSRWLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916  
Qy 901 QLELPVKYAVVMVTSYKYNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 960  
Db 917 QLELPVKYAVVMVTSYKYNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 976  
Qy 961 RLNQTVIWRDPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036  
Qy 1021 PFGIOEEFNATLKGNSLSDWYKTSYKYNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 1080  
Db 1037 PFGIOEEFNATLKGNSLSDWYKTSYKYNFTASENTSRVQHQYQVSNLQGRSLPISLVFLVPV 1096  
Qy 1081 KVEPEVNPPLIIVGSSVGGLLLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1137  
Db 1097 KVEPEVNPPLIIVGSSVGGLLLALITAAALYKLGFFKRYQKMMSEGGPPGABPQ 1153

US-09-902-481A-5

```
Query Match          99.4%; Score 5851; DB 10; Length 1137;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60

Qy 61 RLOVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCTSENTYVKGCLFGLFSGNLROQPOK 120
Db 61 RLOVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCTSENTYVKGCLFGLFSGNLROQPOK 120

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMEKWTVMEOQLKSKTLFSLMOYSEEF 180
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMEKWTVMEOQLKSKTLFSLMOYSEEF 180

Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGVDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGVDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy 301 NFEALKTIONLREKIFAIEGTOTGSSSSFEHMSQEGFSAAITNSGPELLSTVGSYDWAG 360
Db 301 NFEALKTIONLREKIFAIEGTOTGSSSSFEHMSQEGFSAAITNSGPELLSTVGSYDWAG 360

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLGAPRYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480

Qy 481 PRGORARWOCDAVLYGEOGQPMGRFGCAALTVLGDVNGDKLTOVAIGAPGEEDNREGAVYLF 540
Db 481 PRGORARWOCDAVLYGEOGQPMGRFGCAALTVLGDVNGDKLTOVAIGAPGEEDNREGAVYLF 540
```

US-09-902-481A-6

```
Query Match          99.3%; Score 5845; DB 10; Length 1137;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60
Db 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60

Qy 61 RLOVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCTSENTYVKGCLFGLFSGNLROQPOK 120
Db 61 RLOVPVEAVNMSLGLSLAATSPOLLACGPTVHQTCTSENTYVKGCLFGLFSGNLROQPOK 120

Qy 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMEKWTVMEOQLKSKTLFSLMOYSEEF 180
Db 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRRMEKWTVMEOQLKSKTLFSLMOYSEEF 180

Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240
Db 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFL 240

Qy 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGVDAFRSEKSRQELNTVASKPPRDHVFQIN 300
Db 241 TDGEKFGDPLGYEDVPEADREGVIRVYVGVDAFRSEKSRQELNTVASKPPRDHVFQIN 300

Qy 301 NFEALKTIONLREKIFAIEGTOTGSSSSFEHMSQEGFSAAITNSGPELLSTVGSYDWAG 360
Db 301 NFEALKTIONLREKIFAIEGTOTGSSSSFEHMSQEGFSAAITNSGPELLSTVGSYDWAG 360

Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILNRVQSLVLGAPRYQHIGLVAMFR 420

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480
Db 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480

Qy 481 PRGORARWOCDAVLYGEOGQPMGRFGCAALTVLGDVNGDKLTOVAIGAPGEEDNREGAVYLF 540
Db 481 PRGORARWOCDAVLYGEOGQPMGRFGCAALTVLGDVNGDKLTOVAIGAPGEEDNREGAVYLF 540
```

```
Db 481 PRGQARQCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Qy 541 HGTSGSISPHSQRIAGSKLSPRLQYFGQSISGGDLTMDGLVDLTWGAQGHVLLRSQ 600
Db 541 HGTSGSISPHSQRIAGSKLSPRLQYFGQSISGGDLTMDGLVDLTWGAQGHVLLRSQ 600
Qy 601 PVLRVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 601 PVLRVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Qy 661 YDLALDGRPHSRAVFNFTKNSRQTQVLGLTQTCETLKLQLPNCIEDPVPVIRLNF 720
Db 661 YDLALDGRPHSRAVFNFTKNSRQTQVLGLTQTCETLKLQLPNCIEDPVPVIRLNF 720
Qy 721 SLVGTPLSAFGLNRLPVLAEADAQRLFTALPPEKKNCGNDNICODDLSITFSFMSLCLVVG 780
Db 721 SLVGTPLSAFGLNRLPVLAEADAQRLFTALPPEKKNCGNDNICODDLSITFSFMSLCLVVG 780
Qy 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
Db 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
Qy 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPTNKTEF 900
Db 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPTNKTEF 900
Qy 901 QLELPVKYAVVMVTSKYNFTASENTSRVMQHOYQVSNIGORSLSPLSLVFLVPV 960
Db 901 QLELPVKYAVVMVTSKYNFTASENTSRVMQHOYQVSNIGORSLSPLSLVFLVPV 960
Qy 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Qy 1021 PFGIOEEFNATLKGNSLSDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQOGAFVRSQTE 1080
Db 1021 PFGIOEEFNATLKGNSLSDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQOGAFVRSQTE 1080
Qy 1081 KVEPEVENPLPLIYGVSSVGGILLALITAAALYKLGFFKROKMDMSEGGPPGABPQ 1137
Db 1081 KVEPEVENPLPLIYGVSSVGGILLALITAAALYKLGFFKROKMDMSEGGPPGABPQ 1137

RESULT 10
US-09-945-265-4
; Sequence 4, Application US/09945265
; Patent No. US20020123614A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy A.
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Lu, Chafen
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A
; TITLE OF INVENTION: DESIRED CONFORMATION AND METHODS FOR PRODUCING SAME
; FILE REFERENCE: CBN-002CP
; CURRENT APPLICATION NUMBER: US/09/945,265
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,700
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1152
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-945-265-4
```

```
Query Match 99.2%; Score 5836.5; DB 9; Length 1152;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;
Qy 1 FNLDTENAMTFQENARFGQSVVQLOGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCEPI 60
Db 17 FNLDTENAMTFQENARFGQSVVQLOGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCEPI 76
```

```
Qy 61 RLQVPVRAVNMVSLGLSLAATTSPPOLLACGPTVHTQTSENTYVVKGLCFPLGNSLRQOPQK 120
Db 77 RLQVPVRAVNMVSLGLSLAATTSPPOLLACGPTVHTQTSENTYVVKGLCFPLGNSLRQOPQK 136
Qy 121 PPEALRGCPQSDSDIAFLIDSGSIIIPHDFRMKEWSTVMQELKKSKTFLSLMOYSEEF 180
Db 137 PPEALRGCPQSDSDIAFLIDSGSIIIPHDFRMKEFVSTVMQELKKSKTFLSLMOYSEEF 196
Qy 181 RIHTTFKFNQNPNSRLIKITQLGRHTATGLRKVRELFNITNGARKNAFKILFL 240
Db 197 RIHTTFKFNQNPNSRLIKITQLGRHTATGLRKVRELFNITNGARKNAFKILVVI 256
Qy 241 TDGKFGDPLCYEDVPEADREGVIRVVGVDGAFRSKSEKQELNTVASKPRDHVFOIN 300
Db 257 TDGKFGDPLCYEDVPEADREGVIRVVGVDGAFRSKSEKQELNTVASKPRDHVFOIN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQGTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQGTSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLAVMFR 420
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLAVMFR 436
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 496
Qy 481 PRGARQCCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRG-RARQCCDAVLVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
Qy 541 HGTSGSISPHSQRIAGSKLSPRLQYFGQSISGGDLTMDGLVDLTWGAQGHVLLRSQ 600
Db 556 HGTSGSISPHSQRIAGSKLSPRLQYFGQSISGGDLTMDGLVDLTWGAQGHVLLRSQ 615
Qy 601 PVLRVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 616 PVLRVKATMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 675
Qy 661 YDLALDGRPHSRAVFNFTKNSRQTQVLGLTQTCETLKLQLPNCIEDPVPVIRLNF 720
Db 676 YDLALDGRPHSRAVFNFTKNSRQTQVLGLTQTCETLKLQLPNCIEDPVPVIRLNF 735
Qy 721 SLVGTPLSAFGLNRLPVLAEADAQRLFTALPPEKKNCGNDNICODDLSITFSFMSLCLVVG 780
Db 736 SLVGTPLSAFGLNRLPVLAEADAQRLFTALPPEKKNCGNDNICODDLSITFSFMSLCLVVG 795
Qy 781 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 840
Db 796 GPREFNVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACESASSTEV 855
Qy 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPTNKTEF 900
Db 856 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKANKVTSNNMPTNKTEF 915
Qy 901 QLELPVKYAVVMVTSKYNFTASENTSRVMQHOYQVSNIGORSLSPLSLVFLVPV 960
Db 916 QLELPVKYAVVMVTSKYNFTASENTSRVMQHOYQVSNIGORSLSPLSLVFLVPV 975
Qy 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDDIP 1020
Db 976 RLNQTVIWDPRQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDDIP 1035
Qy 1021 PFGIOEEFNATLKGNSLSDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQOGAFVRSQTE 1080
Db 1036 PFGIOEEFNATLKGNSLSDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPQOGAFVRSQTE 1095
Qy 1081 KVEPEVENPLPLIYGVSSVGGILLALITAAALYKLGFFKROKMDMSEGGPPGABPQ 1137
Db 1096 KVEPEVENPLPLIYGVSSVGGILLALITAAALYKLGFFKROKMDMSEGGPPGABPQ 1152
```





121 FPEALRGCPQEDSDIAFLIDGSGSII PHDFRMEKWTSTVMEQLKSKTSLFSLMOYSEEF 180  
138 LPVSRQECPRQEQDIVFLIDGSGSISRNFMNFVRAVISQFORPSTQFSLMQFSNKF 197  
181 RIHTFFKFNPNPRSLIKPITQLLGRTHGTATGLRKVVRELFNTNGARKNAFKILFL 240  
198 QTHTFEFRRTSNPLSLASVHQLQGFTYTATQNVVHRLFHASYGAREDAIKILIVI 257  
241 TDGKFGDPLGYEDVIPELDREGVIRYVIGVDAPRSEKSKQELNTVASKPPRDHVFOIN 300  
258 TDGKKEGSLDYKDVIPMADAAGIIRYAIYGLVAFQNRNSWKELNDIASKESQEHFKVE 317  
301 NFEALKTIONLREKIFAIEGTQGTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 360  
318 DFDALKDIONQLKEKIFAIEGTETITSSSSFELEMAQEGFSAVFTPDGVLGAVGFTWSG 377  
361 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLAVMFR 420  
378 GAFLYPPNWSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437  
421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYYEQTRGGQSVCP 480  
438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSTDVLVLGAPHYYEQTRGGQSVCP 497  
481 PRGQARQCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVATGAPCEEDNRCAVYLF 540  
498 PRGWR-RWMCDAVLYGEOGHPWGRFGAALTVLGDVNGDKLTDVATGAPCEENRCAVYLF 556  
541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTGAGCHVLLRSQ 600  
557 HGVLPSPISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTGAGCHVLLRSQ 616  
601 PVLRYKATMEPNPREVARNVPCNDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
617 PVLWGVGSMQOFIPABIPRSAFECEQVYVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676  
661 YDLALDSGRPHSRVAVNETKSTRQTOVLGTOTCETLKLQLPNCIEDPVSIVLRNLF 720  
677 LDALAPARLSRAIFQETKNSLSRVRLVGLKAHCENFNLLPSCVEDSVIPIILRLNF 736  
721 SILVGTPLSAFGLNRLPWLAEADAQRLFTALFPFEKNCNDNICODDLSITFSFMSLDCLVVG 780  
737 TLVGLPLAFNRLPWLALAAQRYFTASLPFEKNCADHICQDNLGIFSFPGLKSLLVG 796  
781 GPREFNVTVVRNDEGDSVRYTOVTFPFLDLISYRKVSTLONORSORSWRLACESASTEV 840  
797 SNLELNAEVMVWMDGDSYGTITFSPAGLSYRYVAEGQKQGLRSLHLTC--CSAPVG 854  
841 SGALKSTCSINHPIFPENSEVTFNITPDVDSKASIGNKLLKANVTSENNMPTNKTEF 900  
855 SGTWSTSCRINHLIFRGAQITFLATFDVSPKAVGLDRLLLIANVSENNIPRISKITIF 914  
901 QLELPVKAYVMVVTSHGVSTKYLNFATAS-ENTSRVMQHQYQVSNLQORSLPISLFLVP 959  
915 QLELPVKAYVIVVSSHEQTKYLNFSSEESKESHVAMRYQVNNLQORDLPVSNFWVP 974  
960 VRLNQTIVWRPOVTFSENLSTCHTKERLPSHDFLAEIRKAPVNVNCSIAVCQIQCDI 1019  
975 VELNQEAVMMDVEVSHQPNSLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLFRCDV 1034  
1020 PPFQOEERNATLKNLSPDWIKYKSHNHLVSTAETILFNDPSVFTLLPGQCAFVRSOTE 1079  
1035 PSFSQBELDLTKGNLSPGWVRQILQKKVSVSVVABIIIFDSVYSQLPQCAFNRQIT 1094  
1080 TKVEPFEVFNPLFVIGSVGGLLALITALYKLGFFPKQYKQMMSE 1128  
1095 TVLEKYKVHNPPIVIGSVGGLLALITALYKLGFFPKQYKQMMSE 1143

RESULT 13

US-09-891-943-4

; Sequence 4, Application US/09891943

; Publication No. US20030077278A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; TITLE OF INVENTION: NO. US20030077278A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/891,943  
; PRIOR FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-891-943-4  
  
Query Match 58.4%; Score 3436; DB 10; Length 1163;  
Best Local Similarity 60.6%; Pred. No. 5.9e-284;  
Matches 684; Conservative 148; Mismatches 291; Indels 6; Gaps 4;  
  
QY 1 FNLDTENAMTFOENARGFGQSVQVQGSRRVVGAPQOEIVAAANQGRSLVQCQDYSTGSCBPI 60  
DB 20 FNLDTTELTAFRVDAGFGDSVQVYVANSWVVGAPQKIIAANQIGLGYCQGYSTGACBPI 79  
QY 61 RLOYPVEAVNNSGLSLAATTPPOLLACGTPVHQTCSNTYVVKLCFLFGSNLRQQQOK 120  
DB 80 GLQVPPPEAVNNSGLSLAATTPPOLLACGTPVHQTCSNTYVVKLCFLFGSNLRQQQOK 137  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSII PHDFRMEKWTSTVMEQLKSKTSLFSLMOYSEEF 180  
DB 138 LPVSRQECPRQEQDIVFLIDGSGSISRNFMNFVRAVISQFORPSTQFSLMQFSNKF 197  
QY 181 RIHTFFKFNPNPRSLIKPITQLLGRTHGTATGLRKVVRELFNTNGARKNAFKILFL 240  
DB 198 QTHTFEFRRTSNPLSLASVHQLQGFTYTATQNVVHRLFHASYGAREDAIKILIVI 257  
QY 241 TDGKFGDPLGYEDVIPELDREGVIRYVIGVDAPRSEKSKQELNTVASKPPRDHVFOIN 300  
DB 258 TDGKKEGSLDYKDVIPMADAAGIIRYAIYGLVAFQNRNSWKELNDIASKESQEHFKVE 317  
QY 301 NFEALKTIONLREKIFAIEGTQGTGSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 360  
DB 318 DFDALKDIONQLKEKIFAIEGTETITSSSSFELEMAQEGFSAVFTPDGVLGAVGFTWSG 377  
QY 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLAVMFR 420  
DB 378 GAFLYPPNWSPTFINMSQENVMDRDSYLGSTELALWKGVQSLVLGAPRYQHIGKAVIFI 437  
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYYEQTRGGQSVCP 480  
DB 438 QVSRQWRMKAIEVIGTQIGSYFGASLCSVDVDTGSTDVLVLGAPHYYEQTRGGQSVCP 497  
QY 481 PRGQARQCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVATGAPCEEDNRCAVYLF 540  
DB 498 PRGWR-RWMCDAVLYGEOGHPWGRFGAALTVLGDVNGDKLTDVATGAPCEENRCAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTGAGCHVLLRSQ 600  
DB 557 HGVLPSPISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTGAGCHVLLRSQ 616  
QY 601 PVLRYKATMEPNPREVARNVPCNDQVYKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLWGVGSMQOFIPABIPRSAFECEQVYVSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT 676  
QY 661 YDLALDSGRPHSRVAVNETKSTRQTOVLGTOTCETLKLQLPNCIEDPVSIVLRNLF 720

Db 677 LDALAFGLSPRAIFQETKMRSLRVRLGLKAKHCENFNLLLSCEVEDSVIPILRLNF 736  
Qy 721 SLVGTPLSAGNLRLPVLAEDQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
Db 737 TLVGKPLLAFLNRLPMLAALAQRYFTASLPEKNCAGADHICQDNLGIGSFSPGLKSLVVG 796  
Qy 781 GPREFNVTVRNDGDSYRQVTFPPPLDLSYKVSITLQNRQSRWRLACESASSTEV 840  
Db 797 SNLENAEVMWNGEDSYGTTITFSPAGUSYRYVAEGQKQGLRSJLHJTC--CSAPVG 854  
Qy 841 SGALKSTCSINHPITFENSEVTNITFDVDSKASGLNKLKLLKANVTSENNMPTNKTEF 900  
Db 855 SQGTWSTSCRINHLIPRGAQITELATFDVSPKAVGLDRLLLIANVSENNIPTSKTIP 914  
Qy 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTAS--ENTSRVMOHQVQVSNLQORSIPISLVFLVP 959  
Db 915 QLELPVKYAVYVIVVSHHEQFTKYLNFSSEBESHVAMHRYQVNNLQGRDLFVSNFVWP 974  
Qy 960 VRLNQTVIMDRPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDI 1019  
Db 975 VELNQEAVMDVEVSHQFNSLRCSSEKIAPPASDFLAHQKNPVLQCSITAGCLURPRCDV 1034  
Qy 1020 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGQGAFRVSQTE 1079  
Db 1035 PSFSVQEEELFTLKGNLISFGWVRQILQKYSVSVVAEIIPTDSYVQLPGQEAQFMAQTI 1094  
Qy 1080 TKVEPEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMWSE 1128  
Db 1095 TVLEKYRVHNPILPIVINGSSIGLLLLALITAVLYKVGFFKRYKEMMEE 1143

## RESULT 14

US-09-350-259-2  
; Sequence 2, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20020062008A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1161  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-2

Query Match 57.8%; Score 3401; DB 9; Length 1161;  
Best Local Similarity 59.3%; Pred. No. 5.8e-281;  
Matches 669; Conservative 167; Mismatches 285; Indels 8; Gaps 6;

Qy 1 FNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQEIIVAAQNGSLYQCDYSTGSGEPI 60  
Db 17 FNLDVEPTTFQEDAGFGQSVVQFGSRLVVGAPLEVAQNGRLYDCAATGMCQPI 76  
Qy 61 RLQVPEAVNMSLGLSLAATSPQLLACGFTVHQTCSNTYVYKGLCFPLGCSNLRQOPQK 120  
Db 77 FLHIRPEAVNMSLGLTLAASTNGSRLACGFTLHRVCGENSYSKGSCLLLGSRW-EIIQT 135

## RESULT 15

US-09-891-943-2  
; Sequence 2, Application US/09891943  
; Publication No. US20030077278A1

Qy 121 FPEALRGCPQEDSDIAELIDSGSIIHDFRRMKEWSTYMEQLKKSKTLPSIMYOSEEF 180  
Db 136 VPDAITPCPQHEMDIVFLIDSGSIDQNDENQMKGFVOAVNGQFEGTDTTUFALMQYNLL 195  
Qy 181 RIHFTFFKEFQNNPNRSLIKPIITOLLGRTHATGLRNVVRELEFNITNGARKNAFKILFL 240  
Db 196 KIHFTTFQFTSPSQSLVDPIVQLKGLTFTATGILLVTVTLQFHHKNGARKSAKKILIVI 255  
Qy 241 TDEKFGDPLGEDYDIPELDRBGVIRVIVGVDAFRSEKSRQELNVTASPKPRDRHVQIN 300  
Db 256 TDGQKYKDPLEYSVDIPEAKAGIIRYAIVGHAFOGTARQELNLTSSAPPQDHVEKVD 315  
Qy 301 NEERALKTIQOLREKIFAIETGTTGSSSSSEHEMSQEGFSAATISNGPLISTVCSYDAG 360  
Db 316 NFRAALGSIQOLOEKIYAVEGTQSRASSSQHEMSQEGFSTALTMDFLFGAVGSFWSG 375  
Qy 361 GVFLYTSKEKSTFTNMTRVSDMNDAYLGVAAAIIILNRNVQSLVGLGAPRYQHICLVAMFR 420  
Db 376 GAFILYPENMPTTFINMSQENVDMEDSVLGYSTELALWKGVQNLVGLGAPRYQHTKAVFT 435  
Qy 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEYQTRGGQVSVCP 480  
Db 436 QVSRQWRKKAETGVTQIGSYFGASLCSVDVDSNGSTDLLILIGAPHYYEYQTRGGQVSVCP 495  
Qy 481 PRGORARWOCDAVLYGEOGQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 540  
Db 496 PRGORVOMQCDADVLRGEQGHGPRFGAALTVLGDVNBEDKLDIDVAIGAPGOENRGAYLYF 555  
Qy 541 HGTSGSGISPSHQSRLAGSKLSPRLQVFGQSLGGQDLTMDGLVDLTGVGAQGHVLLRSQ 600  
Db 556 HGASEGSIPSHQSRLTASSQLSPRLQVFGQALSGGQDLTQDGLMDLAVGARGQVLLRSL 615  
Qy 601 PVLRVKAIMEFNPREVARNVFECDQVVKGEKAGEVVRVCLHVQKSTRDRREBQIQSVVT 660  
Db 616 PVLKVGVMARFSPVEVAKAYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673  
Qy 661 YDIALDSGRPHSRAVENETKSTRQVTLGLTCTCETLKLQPCNCTEDVPSVILRLNF 720  
Db 674 FDLALDPGRUTSRAIFNETKNPTLTRKKTGLGTHCETLKLKLLPDCVEDVVSPIILHNF 733  
Qy 721 SLVGTPLSAGNLRLPVLAEDQRLFTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
Db 734 SLVREPIPSQNLRLPVLAGSQDLFTASLPEKNCQDGLCEGDLGVTLSFGSLQITLVG 793  
Qy 781 GPREFNVTVRNDGDSYRQVTFPPPLDLSYKVSITLQNRQSRWRLACESASSTEV 840  
Db 794 SSLELNVITVMNAGDSYGTVSVSLYPAGLSHRRVSGAQKQPHQSALRLACETV-PTED 852  
Qy 841 SGALKSTCSINHPITFENSEVTNITFDVDSKASGLNKLKLLKANVTSENNMPTNKTEF 900  
Db 853 EG--LRSSRCVSNHPIFHEGNSNGTFITFDVSVYKATLGDRLMLRASSENKKNKASSKATF 911  
Qy 901 QLELPVKYAVYVMVTVSHGVSTKYLNFTAS--ENTSRVMOHQVQVSNLQORSIPISLVFLVP 959  
Db 912 QLELPVKYAVYVIVVSHHEQFTKYLNFSSEBESHVAMHRYQVNNLQGRDLFVSNFVWP 971  
Qy 960 VRLNQTVIMDRPQVTFSENISSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDI 1019  
Db 972 VLLNGVAVMDVVMVMEAPSQSL--PCVSEKPKPPQSHSDFLTQISRSPMLDCSIADCLQPRCDV 1029  
Qy 1020 PFGIOEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGQGAFRVSQTE 1079  
Db 1030 PSFSVQEEELFTLKGNLISFGWVRQILQKYSVSVVAEIIPTDSYVQLPGQEAQFMAQME 1089  
Qy 1080 TKVEPEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMWSE 1128  
Db 1090 MVLEEDBVNAPIIMGSSVGCALLLALITATLYKLGFFKRYKEMLEED 1138



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds  
(without alignments)  
3278.416 Million cell updates/sec

Title: US-09-902-481B-4  
Perfect score: 5884  
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5852	99.5	1153	1 US-08-173-497-3	Sequence 3, Appli
2	5852	99.5	1153	1 US-08-286-889-3	Sequence 3, Appli
3	5852	99.5	1153	1 US-08-485-618-3	Sequence 3, Appli
4	5852	99.5	1153	1 US-08-362-652-3	Sequence 3, Appli
5	5852	99.5	1153	2 US-08-605-672-3	Sequence 3, Appli
6	5852	99.5	1153	2 US-08-482-293A-3	Sequence 3, Appli
7	5852	99.5	1153	2 US-08-943-363-3	Sequence 3, Appli
8	5852	99.5	1153	3 US-09-193-043-3	Sequence 3, Appli
9	5852	99.5	1153	4 US-09-688-307A-3	Sequence 3, Appli
10	5852	99.5	1153	4 US-09-350-259-3	Sequence 3, Appli
11	5821.5	98.9	1152	2 US-08-476-062A-43	Sequence 43, Appl
12	5821.5	98.9	1152	6 PCT-US96-01314-43	Sequence 43, Appl
13	5821.5	98.9	1152	6 5424399-2	Patent No. 5424399
14	3459	58.8	1163	5 US-08-476-062A-44	Sequence 44, Appl
15	3459	58.8	1163	5 PCT-US96-01314-44	Sequence 44, Appl
16	3436	58.4	1163	1 US-08-173-497-4	Sequence 4, Appli
17	3436	58.4	1163	1 US-08-286-889-4	Sequence 4, Appli
18	3436	58.4	1163	1 US-08-485-618-4	Sequence 4, Appli
19	3436	58.4	1163	1 US-08-362-652-4	Sequence 4, Appli
20	3436	58.4	1163	2 US-08-605-672-4	Sequence 4, Appli
21	3436	58.4	1163	2 US-08-482-293A-4	Sequence 4, Appli
22	3436	58.4	1163	2 US-08-943-363-4	Sequence 4, Appli
23	3436	58.4	1163	3 US-09-193-043-4	Sequence 4, Appli
24	3436	58.4	1163	4 US-09-688-307A-4	Sequence 4, Appli
25	3436	58.4	1163	4 US-09-350-259-4	Sequence 4, Appli
26	3401	57.8	1161	1 US-08-173-497-2	Sequence 2, Appli
27	3401	57.8	1161	1 US-08-286-889-2	Sequence 2, Appli

28	3401	57.8	1161	1 US-08-485-618-2	Sequence 2, Appli
29	3401	57.8	1161	1 US-08-362-652-2	Sequence 2, Appli
30	3401	57.8	1161	2 US-08-605-672-2	Sequence 2, Appli
31	3401	57.8	1161	2 US-08-482-293A-2	Sequence 2, Appli
32	3401	57.8	1161	2 US-08-943-363-2	Sequence 2, Appli
33	3401	57.8	1161	3 US-09-193-043-2	Sequence 2, Appli
34	3401	57.8	1161	4 US-09-688-307A-2	Sequence 2, Appli
35	3401	57.8	1161	4 US-09-350-259-2	Sequence 2, Appli
36	3385.5	57.5	1161	1 US-08-485-618-99	Sequence 99, Appl
37	3385.5	57.5	1161	2 US-08-605-672-99	Sequence 99, Appl
38	3385.5	57.5	1161	2 US-08-482-293A-99	Sequence 99, Appl
39	3385.5	57.5	1161	2 US-08-943-363-99	Sequence 99, Appl
40	3385.5	57.5	1161	3 US-09-193-043-99	Sequence 99, Appl
41	3385.5	57.5	1161	4 US-09-688-307A-99	Sequence 99, Appl
42	3385.5	57.5	1161	4 US-09-350-259-99	Sequence 99, Appl
43	3224.5	54.8	1161	3 US-09-193-043-55	Sequence 55, Appl
44	3224.5	54.8	1161	4 US-09-688-307A-55	Sequence 55, Appl
45	3224.5	54.8	1161	4 US-09-350-259-55	Sequence 55, Appl

ALIGNMENTS

RESULT 1  
US-08-173-497-3  
; Sequence 3, Application US/08173497  
; Patent No. 5437958  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van Der Vieren, Monica  
; TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha  
; TITLE OF INVENTION: Subunit  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 S. Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/173,497  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5437958and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/31363  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-173-497-3

Query Match 99.5% Score 5852; DB 1; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGRSVVVGAPQEIIVANQRGSLYQCDYSTGSGCEPI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGRSVVVGAPQEIIVANQRGSLYQCDYSTGSGCEPI 76

```
QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLGSLNLRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLGSLNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPRSLIKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPRSLIKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILVVI 256
QY 241 TDGEKFGPLGVEDVPELDRGVIRYVIGDAPRSEKSRQELNVTASKPPDRHVQIN 300
Db 257 TDGEKFGPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNVTASKPPDRHVQIN 316
QY 301 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 360
Db 317 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 420
Db 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGYAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 436
QY 421 QVNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVETRGGOVSVCLP 480
Db 437 QVNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVETRGGOVSVCLP 496
QY 481 PRGQARWOCDAVLVYGEQOQWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVLYF 540
Db 497 PRGQARWOCDAVLVYGEQOQWGRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVLYF 556
QY 541 HCTSGSGTSPSHSQRISQISKLSPRIQYFGQSLSGGQDLTMDGLVLTIVGAQGHVLLRSQ 600
Db 557 HCTSGSGTSPSHSQRISQISKLSPRIQYFGQSLSGGQDLTMDGLVLTIVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMEFNPVARNVPCNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660
Db 617 PVLRVKAIMEFNPVARNVPCNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676
QY 661 YDLALDGRPHSRVAFNETKNSRRQTQVLGLTQTCETILKQLPNCIEDPVPPIVRLNF 720
Db 677 YDLALDGRPHSRVAFNETKNSRRQTQVLGLTQTCETILKQLPNCIEDPVPPIVRLNF 736
QY 721 SLVGTPLSAFGLNLRVLAEDAQRLFTALPPFKKNCNDNI CODDLSITFSFMSLCLVVG 780
Db 737 SLVGTPLSAFGLNLRVLAEDAQRLFTALPPFKKNCNDNI CODDLSITFSFMSLCLVVG 796
QY 781 GPREFNVTVTVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSWRKACESASSTEV 840
Db 797 GPREFNVTVTVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSWRKACESASSTEV 856
QY 841 SCALKSTSCSINHPIFPENSEVTFNI TEDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
Db 857 SCALKSTSCSINHPIFPENSEVTFNI TEDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVTVSHGVSTKYNFTASENTRVWQHQQYQVSNLQSRSLPISLFLVLPV 960
Db 917 QLELPVKYAVVMVTVSHGVSTKYNFTASENTRVWQHQQYQVSNLQSRSLPISLFLVLPV 976
QY 961 RLNQTVIWDPRQVTFSENLSSSTCHTKERLPSSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
Db 977 RLNQTVIWDPRQVTFSENLSSSTCHTKERLPSSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1036
QY 1021 PFGIQEEFNATLKGNSFDWYIKTSHNHLILVSTAEIILFNDVSFTLLPGQGFVRSQTET 1080
Db 1037 PFGIQEEFNATLKGNSFDWYIKTSHNHLILVSTAEIILFNDVSFTLLPGQGFVRSQTET 1096
QY 1081 KYEPFVVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKQYKQMMSEGGPPGABPQ 1137
Db 1097 KYEPFVVPNPLPLIVGSSVGGLLLLALITAALYKLGFFKQYKQMMSEGGPPGABPQ 1153
```

```
RESULT 2
US-08-286-889-3
; Sequence 3, Application US/08286889
; Patent No. 5470953
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Mich
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/286.889
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: P38,659
; REFERENCE/DOCKET NUMBER: 27866/32168
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1153 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-286-889-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVOLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARGFGQSVVOLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLGSLNLRQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVKGCLFGLGSLNLRQPOK 136
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKSKTLFSLMOYSEEF 180
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKWVSTVMEQLKSKTLFSLMOYSEEF 196
QY 181 RIHFTFKFQNNPRSLIKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILFLL 240
Db 197 RIHFTFKFQNNPRSLIKPTQLLGRTHATGIRKVVRELFNTNGARKNAFKILVVI 256
QY 241 TDGEKFGPLGVEDVPELDRGVIRYVIGDAPRSEKSRQELNVTASKPPDRHVQIN 300
Db 257 TDGEKFGPLGVEDVPEADREGVIRYVIGDAPRSEKSRQELNVTASKPPDRHVQIN 316
QY 301 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 360
Db 317 NFEALKTIONQUREKIFAIEGTQGTSSSSFEHMSQEGFSAAITSNGLLSTVGSYDWAG 376
```

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP 480  
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP 496  
QY 481 PRGORARWQCDVLYGEGQGWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYLYF 540  
Db 497 PRGORARWQCDVLYGEGQGWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYLYF 556  
QY 541 HGTSGSGISPSHQRINGSKSLRPLQYFGQSLSGQDLTMDGLVDLTGAGQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHQRINGSKSLRPLQYFGQSLSGQDLTMDGLVDLTGAGQGHVLLRSQ 616  
QY 601 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRRREGQIQSVVT 660  
Db 617 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRRREGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 720  
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPVILRLNF 736  
QY 721 SLVGTPLSARGNLRPLVAEDAORLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSARGNLRPLVAEDAORLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRSLACASASTEV 840  
Db 797 GPREFNVTVVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRSLACASASTEV 856  
QY 841 SGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Db 857 SGALKSTCSINHPFIPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMQHOVQVSNLQORSJPLSLVFLVPV 960  
Db 917 QLELPVKYAVYVMTSHGVSTKYNFTASENTSRVMQHOVQVSNLQORSJPLSLVFLVPV 976  
QY 961 RLNOTVTDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
Db 977 RLNOTVTDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036  
QY 1021 PFGIOEBFNATLKNLSFDWYIKTSHNELLIVSTAEILFNDSVFTLLPGQAGFVRSQTE 1080  
Db 1037 PFGIOEBFNATLKNLSFDWYIKTSHNELLIVSTAEILFNDSVFTLLPGQAGFVRSQTE 1096  
QY 1081 KVEPFEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFKKQYKDMSEGPPGAEPQ 1137  
Db 1097 KVEPFEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFKKQYKDMSEGPPGAEPQ 1153

## RESULT 3

US-08-485-618-3  
; Sequence 3, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,618  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32797  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; SEQUENCE CHARACTERISTICS:  
; INFORMATION FOR SEQ ID NO: 3:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-485-618-3

Query Match 99.5%; Score 5852; DB 1; Length 1153;

Best Local Similarity 99.2%; Pred. No. 0;

Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY	1	FNLDTENAMTFQENARGFGOSVVQLOGSRVVVGAPOEIVAAORGSILYQCDYSTGSCPEI	60
Db	17	FNLDTENAMTFQENARGFGOSVVQLOGSRVVVGAPOEIVAAORGSILYQCDYSTGSCPEI	76
QY	61	RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGKLCFLFGSNLRQOPQK	120
Db	77	RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGKLCFLFGSNLRQOPQK	136
QY	121	FPALRGCPQEDSDIAPLIDGSGSIIPHDFRMRKEWVSTWMEQLKKSKTLFSLMQYSEEF	180
Db	137	FPALRGCPQEDSDIAPLIDGSGSIIPHDFRMRKEFVSTWMEQLKKSKTLFSLMQYSEEF	196
QY	181	RIHFTKFEFQNNPNRSLIKPITQLLGRTHATGLRKVRVRELNITNGARKNAFKILFL	240
Db	197	RIHFTKFEFQNNPNRSLIKPITQLLGRTHATGLRKVRVRELNITNGARKNAFKILVVI	256
QY	241	TDGEKFGDPLGVEDVPELDRGVIRYVIGVDAFRSEKSRQELNITVASKPPRDHVFQIN	300
Db	257	TDGEKFGDPLGVEDVPEADREGVIRYVIGVDAFRSEKSRQELNITASKPPRDHVFQIN	316
QY	301	NFEALKTIONLREKIFAIEGTQTGSSSSSEHEMSQGFSAAITNSGNPLSTVGSYDWAG	360
Db	317	NFEALKTIONLREKIFAIEGTQTGSSSSSEHEMSQGFSAAITNSGNPLSTVGSYDWAG	376
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRYQHIGLVAMFR	420
Db	377	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRYQHIGLVAMFR	436
QY	421	QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP	480
Db	437	QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVCP	496
QY	481	PRGORARWQCDVLYGEGQGWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYLYF	540
Db	497	PRGORARWQCDVLYGEGQGWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYLYF	556
QY	541	HGTSGSGISPSHQRINGSKSLRPLQYFGQSLSGQDLTMDGLVDLTGAGQGHVLLRSQ	600

Db 557 HGTSGSGISPSHQIAGSKLSPRLQYFGQSLGGDLTMDGLVDLTGVAQGHVLLRSQ 616  
QY 601 PVLRVKAIEMFNPVARNVFNCDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
Db 617 PVLRVKAIEMFNPVARNVFNCDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676  
QY 661 YDLALDSDGRPHRAVFNKSTRROTQVLGTQTCETLKLQLPNCIEDPVPSPVLRNLF 720  
Db 677 YDLALDSDGRPHRAVFNKSTRROTQVLGTQTCETLKLQLPNCIEDPVPSPVLRNLF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPFKNGNDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPFKNGNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GREENVTVVRNDEGDSVRTQVTFPFLDLSYRKVSTLQNSQSRWSLACESASSTEV 840  
Db 797 GREENVTVVRNDEGDSVRTQVTFPFLDLSYRKVSTLQNSQSRWSLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIPFENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKAYVMVTVSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQORSPLPSLVPLVPV 960  
Db 917 QLELPVKAYVMVTVSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQORSPLPSLVPLVPV 976  
QY 961 RLNOTVWRPQVTFSENLSSTCHTKERLPSPHSDFLAELRKAPVNCSTAVCQRIQCDIP 1020  
Db 977 RLNOTVWRPQVTFSENLSSTCHTKERLPSPHSDFLAELRKAPVNCSTAVCQRIQCDIP 1036  
QY 1021 FFGIOEFNATLKGNSLSPDWYIKTSHNHLIIYSTAELLFNDVSFTLLPCQAGFVRSQET 1080  
Db 1037 FFGIOEFNATLKGNSLSPDWYIKTSHNHLIIYSTAELLFNDVSFTLLPCQAGFVRSQET 1096  
QY 1081 KVEPEVPNPLPIVGVSSVGGLLLLALITAAIYKLGFFKQYKQDMSEGPPGABPQ 1137  
Db 1097 KVEPEVPNPLPIVGVSSVGGLLLLALITAAIYKLGFFKQYKQDMSEGPPGABPQ 1153

RESULT 4  
US-08-362-652-3  
; Sequence 3, Application US/08362652  
; Patent No. 5766850  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,652  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.

; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32391  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEFAX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-362-652-3  
  
Query Match 99.5%; Score 5852; DB 1; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIVAANQORGLYOCYSTGSCPEI 60  
Db 17 FNLDTENAMTFOENARGFGQSVVQLQSGRVVVGAPQEIVAANQORGLYOCYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLAGCPTVHOTCSENTYVKGICFLFGSNLRQPOK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLAGCPTVHOTCSENTYVKGICFLFGSNLRQPOK 136  
QY 121 PFEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKKQVSTVMQKKSKTLFSLMQYSEFP 180  
Db 137 PFEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKKQVSTVMQKKSKTLFSLMQYSEFP 196  
QY 181 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 240  
Db 197 RHFTFKFQNNPNRSLIKPITQLLGRTHATGLRKVVRELFNITNGARKNAFKILFLL 256  
QY 241 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSRQELNNTASKPRDRHVQFN 300  
Db 257 TDGEKFGDPLGYEDVIPELDREGVIRYVIGVDAPFRSEKSRQELNNTASKPRDRHVQFN 316  
QY 301 NFEALKTQNLREKIPALBGTQTCSSSSFEHMSQEGFSAITNSGPLLSTVSGYDWAQ 360  
Db 317 NFEALKTQNLREKIPALBGTQTCSSSSFEHMSQEGFSAITNSGPLLSTVSGYDWAQ 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAIILRNKQSLVLGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAIILRNKQSLVLGAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 480  
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCP 496  
QY 481 PRGQARWQCDALYGEQGPWRFGAALTVLVDVNGDKLTDAIAGPEENRGAIVLF 540  
Db 497 PRGQARWQCDALYGEQGPWRFGAALTVLVDVNGDKLTDAIAGPEENRGAIVLF 556  
QY 541 HGTSGSGISPSHQIAGSKLSPRLQYFGQSLGGDLTMDGLVDLTGVAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHQIAGSKLSPRLQYFGQSLGGDLTMDGLVDLTGVAQGHVLLRSQ 616  
QY 601 PVLRVKAIEMFNPVARNVFNCDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660  
Db 617 PVLRVKAIEMFNPVARNVFNCDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT 676  
QY 661 YDLALDSDGRPHRAVFNKSTRROTQVLGTQTCETLKLQLPNCIEDPVPSPVLRNLF 720  
Db 677 YDLALDSDGRPHRAVFNKSTRROTQVLGTQTCETLKLQLPNCIEDPVPSPVLRNLF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPFKNGNDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFGLNRPVLAEDAQLFTALPFPFKNGNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GREENVTVVRNDEGDSVRTQVTFPFLDLSYRKVSTLQNSQSRWSLACESASSTEV 840



Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPPIFENSEVTFNITFDVDSKASLGNKULLKANVTSENMMPTNKTEF 900  
Db 857 SGALKSTSCSINHPPIFENSEVTFNITFDVDSKASLGNKULLKANVTSENMMPTNKTEF 916  
Qy 901 QLELPVKYAYVMVTVSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLISLVLVPV 960  
Db 917 QLELPVKYAYVMVTVSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLISLVLVPV 976  
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036  
Qy 1021 FPGIOEFNATLKGNSLFDWIKTSHNHLIIVSTAEILFNDVSTLLPGQGAFFVRSQTET 1080  
Db 1037 FPGIOEFNATLKGNSLFDWIKTSHNHLIIVSTAEILFNDVSTLLPGQGAFFVRSQTET 1096  
Qy 1081 KVEPFEVPNPLPLTVGSSVGLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1137  
Db 1097 KVEPFEVPNPLPLTVGSSVGLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1153

RESULT 5

US-08-605-672-3  
; Sequence 3, Application US/08605672  
; Patent No. 5817515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vleren, Monica  
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/605,672  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-605-672-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAANQSGSIYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVGAPOEIVAAANQSGSIYQCDYSTGSCPEI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGICFLPGSNLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGICFLPGSNLRQOPQK 136  
Qy 121 FPEALGCCPOEDSDIAFLIDGSGSIIPHDFRRMKEWSTVMEOLKSKTKLFSLMQYSEEF 180  
Db 137 FPEALGCCPOEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEOLKSKTKLFSLMQYSEEF 196  
Qy 181 RIHFTFKEFQNNPNRSLIKPITQLLGRTHATGLARKVRELFNITNGARKNAFKILFL 240  
Db 197 RIHFTFKEFQNNPNRSLVKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILV 256  
Qy 241 TDGEKFGDPLGYEDVPELDREGVIRYVGVGDAFRSEKSRQBELNVTASKPPRHDHVFQIN 300  
Db 257 TDGEKFGDPLGYEDVPELDREGVIRYVGVGDAFRSEKSRQBELNVTASKPPRHDHVFQIN 316  
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSEFHEMSQEGFSAAITSGNPLLSVTGSDWAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSEFHEMSQEGFSAAITSGNPLLSVTGSDWAG 376  
Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVGLAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVGLAPRYQHIGLVAMFR 436  
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVSCPL 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVSCPL 496  
Qy 481 PRGORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYL 540  
Db 497 PRGORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYL 556  
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 616  
Qy 601 PVLRVKAIMEFNPREVARNVFECDNQVVKGKEAGEVRVCLHVQKSTDRDLREGQIQSVVT 660  
Db 617 PVLRVKAIMEFNPREVARNVFECDNQVVKGKEAGEVRVCLHVQKSTDRDLREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRAVENETKNSRTOFVLGLTCTETLKLQPNCTEDPVSPVLRLNF 720  
Db 677 YDLALDSGRPHSRAVENETKNSRTOFVLGLTCTETLKLQPNCTEDPVSPVLRLNF 736  
Qy 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPPIFENSEVTFNITFDVDSKASLGNKULLKANVTSENMMPTNKTEF 900  
Db 857 SGALKSTSCSINHPPIFENSEVTFNITFDVDSKASLGNKULLKANVTSENMMPTNKTEF 916  
Qy 901 QLELPVKYAYVMVTVSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLISLVLVPV 960  
Db 917 QLELPVKYAYVMVTVSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLISLVLVPV 976  
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036

QY 1021 PFGIQEENATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQGFVRSQTET 1080  
Db 1037 PFGIQEENATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQGFVRSQTET 1096  
QY 1081 KVEPEFVNPLPIIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGPPGABPQ 1137  
Db 1097 KVEPEFVNPLPIIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGPPGABPQ 1153

## RESULT 6

US-08-482-293A-3  
; Sequence 3, Application US/08482293A  
; Patent No. 5831029  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,293A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: amino acid  
; LENGTH: 1153 amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-482-293A-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;  
QY 1 FNLDTENMTFQENARGFGSVQLOGSRVVVGAPOEIVAAORGLSYQCQDYSGSCPEI 60  
Db 17 FNLDTENMTFQENARGFGSVQLOGSRVVVGAPOEIVAAORGLSYQCQDYSGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFPLFGSNLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVVKGLCFPLFGSNLRQOPQK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKSKTLFSLMQYSEEF 180

Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196  
QY 181 RHITFTKFGQNNPNRSLIKPITQLLGRTHATGKRVVRELFINITNGARKNAFKILFLL 240  
Db 197 RHITFTKFGQNNPNRSLIKPITQLLGRTHATGKRVVRELFINITNGARKNAFKILVVI 256  
QY 241 TDGEKFGDPLGYEDVIPELDRGVIRYVIGVDAERSEKSRQELNVTASKPPDRHVQIN 300  
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAERSEKSRQELNVTASKPPDRHVQIN 316  
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSFHEHMOEGFSAAITNGPLLSVTGVSVDWAG 360  
Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSFHEHMOEGFSAAITNGPLLSVTGVSVDWAG 376  
QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 420  
Db 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVIGAPRYQHIGLVAMPR 436  
QY 421 QNTGMWESNANVKGTOIGAYTCASICSVDDVDSNGSTDLVIGAPHVYEQTRGGQVSVCL 480  
Db 437 QNTGMWESNANVKGTOIGAYTCASICSVDDVDSNGSTDLVIGAPHVYEQTRGGQVSVCL 496  
QY 481 PRGQARWQCDVLYGEOGQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRCGAVILF 540  
Db 497 PRGQARWQCDVLYGEOGQPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEEDNRCGAVILF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTIVGAQHVLRLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTIVGAQHVLRLRSQ 616  
QY 601 PVLRVKAJMEFNPREVARNVFECDQVVKGEAGVRVCLHVQKSTRDLRREGQOSVVT 660  
Db 617 PVLRVKAJMEFNPREVARNVFECDQVVKGEAGVRVCLHVQKSTRDLRREGQOSVVT 676  
QY 661 YDLALDSGRPHSRAVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720  
Db 677 YDLALDSGRPHSRAVFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 736  
QY 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFEPPLDLSYKRVSTLQNCORSORSLACESASSTEV 840  
Db 797 GPREFNVTVTVRNDGEDSYRTQVTFEPPLDLSYKRVSTLQNCORSORSLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYOVSNLQBSLPISLVFLVPV 960  
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTSRVMOHQYOVSNLQBSLPISLVFLVPV 976  
QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVUNCISIAVCORIQCDIP 1020  
Db 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVUNCISIAVCORIQCDIP 1036  
QY 1021 FFGIQEENATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQGFVRSQTET 1080  
Db 1037 FFGIQEENATLKGNSLFDWYIKTSHNHLIIIVSTAEILFNDVSFTLLPGQGFVRSQTET 1096  
QY 1081 KVEPEFVNPLPIIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGPPGABPQ 1137  
Db 1097 KVEPEFVNPLPIIVGSSVGGLLLLALITAAALYKLGFFKRYQKDMSEGPPGABPQ 1153

## RESULT 7

US-08-943-363-3  
; Sequence 3, Application US/08943363  
; Patent No. 5837478  
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943,363  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-943-363-3

Query Match 99.5%; Score 5852; DB 2; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAQNGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSGLSLAATSPQLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 120  
DB 77 RLQVPVEAVNMSGLSLAATSPQLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQSDIAFLIDGSGSIIIPHDPRMKEWSTVMEQKSKTFLSLMOYSEEF 180  
DB 137 FPEALRGCPQSDIAFLIDGSGSIIIPHDPRMKEFVSTVMEQKSKTFLSLMOYSEEF 196

QY 181 RIHFTFKFQNNPNRSLIKPITQLGRTHATGLRKVVRELFINITNGARKNAKILFL 240  
DB 197 RIHFTFKFQNNPNRSLVKEITQLGRTHATGIRKVVRELFINITNGARKNAKILVVI 256

QY 241 TDGKFGDPLGYEDVIPLEDEGVIRYVIGVDAPRSEKSEQLNTVASKPRDHVFQIN 300  
DB 257 TDGKFGDPLGYEDVIPLEDEGVIRYVIGVDAPRSEKSEQLNTIASKPRDHVFQIN 316

QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAG 360  
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPLLTSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVILGAPRYOHIGLVAMER 420  
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVILGAPRYOHIGLVAMER 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVILGAPHYEYTRGQSVCP 480  
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLLVILGAPHYEYTRGQSVCP 496

QY 481 PRGQARWQCDVLYGEGQCPWGRFGAALTIVLDVNGDKLTDVAIGAPCEEDNRGAVYLF 540  
DB 497 PRGQARWQCDVLYGEGQCPWGRFGAALTIVLDVNGDKLTDVAIGAPCEEDNRGAVYLF 556

QY 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600  
DB 557 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

QY 601 PVLRVKAIMFEFNPREVARNVFECNDQVVRKEAGEVVRVCLHVQKSTRDLREGOIQSVVT 660  
DB 617 PVLRVKAIMFEFNPREVARNVFECNDQVVRKEAGEVVRVCLHVQKSTRDLREGOIQSVVT 676

QY 661 YDLALDSGRPHSRVAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720  
DB 677 YDLALDSGRPHSRVAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSIVLRLNF 736

QY 721 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFKXNCGNDNICODDLSITPFSMSLDCLVVG 780  
DB 737 SLVGTPLSAFGNLRPVLAEDAQRLFTALPPFKXNCGNDNICODDLSITPFSMSLDCLVVG 796

QY 781 GPREFNVTVTRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSQRSWRLACESASSTEV 840  
DB 797 GPREFNVTVTRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSQRSWRLACESASSTEV 856

QY 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 900  
DB 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKLLKANVTSENNMPTNKTEF 916

QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSUPISLVFLVPV 960  
DB 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSUPISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDSDFLAELKAPVYVNCIAVCQRIQCDIP 1020  
DB 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSSHDSDFLAELKAPVYVNCIAVCQRIQCDIP 1036

QY 1021 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOTET 1080  
DB 1037 FFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOTET 1096

QY 1081 KVEPPEVPNPLPLTVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1137  
DB 1097 KVEPPEVPNPLPLTVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1153

RESULT 8  
US-09-193-043-3  
; Sequence 3, Application US/09193043  
; Patent No. 6251395  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6251395el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/193,043  
; CURRENT FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-193-043-3

Query Match 99.5%; Score 5852; DB 3; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAQEIIVANQORGLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAQEIIVANQORGLYQCDYSTGSCPEI 76

QY 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVVKGLCFGLFSGNLROQPOK 120  
DB 77 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVVKGLCFGLFSGNLROQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMBOLKSKTLFSLMOYSBEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMBOLKSKTLFSLMOYSBEF 196

QY 181 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGLRKVRELFINITNGARKNAFKILV 240  
DB 197 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGLRKVRELFINITNGARKNAFKILV 256

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRVYVIGDAPRSEKSOELNTVASKPPRDHVFQIN 300  
DB 257 TDGEKFGDPLGYEDVIPEDREGVIRVYVIGDAPRSEKSOELNTVASKPPRDHVFQIN 316

QY 301 NFPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSOEGFSAAITNSGPLLSTVGSYDAG 360  
DB 317 NFPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSOEGFSAAITNSGPLLSTVGSYDAG 376

QY 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVLAGAPRYOHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVLAGAPRYOHIGLVAMFR 436

QY 421 QNTGWESNANVKTQICAYGASLCSVDVDSNGSTDLVLGAPHYETQTRGGQSVCP 480  
DB 437 QNTGWESNANVKTQICAYGASLCSVDVDSNGSTDLVLGAPHYETQTRGGQSVCP 496

QY 481 PRGORARQCDAVLYGEGQPGWGRFAGALTVLGVNGDKLTDVAIGAPGEDNRGAVLYF 540  
DB 497 PRGORARQCDAVLYGEGQPGWGRFAGALTVLGVNGDKLTDVAIGAPGEDNRGAVLYF 556

QY 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
DB 557 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616

QY 601 PVLARKVKAIMEFNPREVARNVECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLARKVKAIMEFNPREVARNVECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676

QY 661 YDLALDSGRPHSRVAFNETKNSRRQTOVLGLTQTCETLKLQLENCIEDPVSIVLRNF 720  
DB 677 YDLALDSGRPHSRVAFNETKNSRRQTOVLGLTQTCETLKLQLENCIEDPVSIVLRNF 736

QY 721 SLVGTPTLSAFGNLRPVLAEADQRLFTALFPEKKNCGNDNICQDDLSTIFSPMSLDCLVVG 780  
DB 737 SLVGTPTLSAFGNLRPVLAEADQRLFTALFPEKKNCGNDNICQDDLSTIFSPMSLDCLVVG 796

QY 781 GPREFNVTVVRNDEGDSYRTQVTFEPFLDLSYRKVSTLQNRORSORWKLACASSTEV 840  
DB 797 GPREFNVTVVRNDEGDSYRTQVTFEPFLDLSYRKVSTLQNRORSORWKLACASSTEV 856

QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
DB 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916

QY 901 QLELPVKYAVYVMTSHGVSTKYNLNTASENTSRVMQHQYQVSNLQORSULPISLVFLVP 960

DB 917 QLELPVKYAVYVMTSHGVSTKYNLNTASENTSRVMQHQYQVSNLQORSULPISLVFLVP 976

QY 961 RINQTVIWDPRQVTFSENLSTCHTKERLPKSHDFLABLRKAPVNVNCISIAVCQRIQCDIP 1020

DB 977 RINQTVIWDPRQVTFSENLSTCHTKERLPKSHDFLABLRKAPVNVNCISIAVCQRIQCDIP 1036

QY 1021 FPGIOEENATLKGNLSFDWYIKTSHNLLIVSTAEILLNDVFTLLPGQGAFFVSQET 1080

DB 1037 FPGIOEENATLKGNLSFDWYIKTSHNLLIVSTAEILLNDVFTLLPGQGAFFVSQET 1096

QY 1081 KVEPEVENPLPLIIVGSSVGGLLILALITAAALYKLGFFKROVKOMWSSGGPPGABPO 1137

DB 1097 KVEPEVENPLPLIIVGSSVGGLLILALITAAALYKLGFFKROVKOMWSSGGPPGABPO 1153

RESULT 9  
US-09-688-307A-3  
; Sequence 3, Application US/09688307A  
; Patent No. 6432404  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; TITLE OF INVENTION: No. 6432404el Human Beta-2  
; FILE REFERENCE: 27866/36646  
; CURRENT APPLICATION NUMBER: US/09/688,307A  
; CURRENT FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/605,672  
; PRIOR FILING DATE: 1996-02-22  
; PRIOR APPLICATION NUMBER: 08/173,497  
; PRIOR FILING DATE: 1993-12-23  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PR1  
; ORGANISM: Homo sapiens  
US-09-688-307A-3

Query Match 99.5%; Score 5852; DB 4; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAQEIIVANQORGLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRRVVVGAQEIIVANQORGLYQCDYSTGSCPEI 76

QY 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVVKGLCFGLFSGNLROQPOK 120  
DB 77 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSENTYVVKGLCFGLFSGNLROQPOK 136

QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWSTVMEQKKSKTLFSLMOYSBEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSBEF 196

QY 181 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGLRKVRELFINITNGARKNAFKILV 240  
DB 197 RIHFTFKFQNNPNRSLIKPITOLLGRTHATGLRKVRELFINITNGARKNAFKILV 256

QY 241 TDGEKFGDPLGYEDVIPEDREGVIRVYVIGDAPRSEKSOELNTVASKPPRDHVFQIN 300  
DB 257 TDGEKFGDPLGYEDVIPEDREGVIRVYVIGDAPRSEKSOELNTVASKPPRDHVFQIN 316

QY 301 NFPEALKTIQNLREKIFAIEGTQTGSSSSFEHMSOEGFSAAITNSGPLLSTVGSYDAG 360

Db 317 NFEALKTIONQRLREKIFAIRGTQTGSSSSFEHEMSQGFSAAITNSGPLLSTVGSYDWAG 376  
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 436  
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSCPL 480  
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSCPL 496  
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540  
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 556  
Qy 541 HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
Qy 601 PVLRVKAIMBFPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 660  
Db 617 PVLRVKAIMBFPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736  
Qy 721 SLVGTPLSAGFNLRPVLAEADAQRLFTALFPPEKNCNDNICQDDLSITFFSMDLCLVVG 780  
Db 737 SLVGTPLSAGFNLRPVLAEADAQRLFTALFPPEKNCNDNICQDDLSITFFSMDLCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916  
Qy 901 QLELPVKYAVMVVTSYGVSTKYNLFTASENWSRWMOHQVQVSNLQORSPLISLVFLVPV 960  
Db 917 QLELPVKYAVMVVTSYGVSTKYNLFTASENWSRWMOHQVQVSNLQORSPLISLVFLVPV 976  
Qy 961 RLNOTVLDROPVTFESNLSTCHTKERLPKSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1020  
Db 977 RLNOTVLDROPVTFESNLSTCHTKERLPKSHSDFLAELRKAPVNVCSIAVCORIQCDIP 1036  
Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDVSFTLLPGQAFVRSQTET 1080  
Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDVSFTLLPGQAFVRSQTET 1096  
Qy 1081 KVEPFEPVNPPLIVGSSVGLLILALITAAALYKLGFFKROYKDMMSGEGPPGAEPQ 1137  
Db 1097 KVEPFEPVNPPLIVGSSVGLLILALITAAALYKLGFFKROYKDMMSGEGPPGAEPQ 1153

RESULT 10  
US-09-350-259-3  
; Sequence 3, Application US/09350259  
; Patent No. 6620915  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6620915el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-350-259-3

Query Match 99.5%; Score 5852; DB 4; Length 1153;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGOSVVOQLGSRVVVVGAPQEIIVAAQORGSILYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGOSVVOQLGSRVVVVGAPQEIIVAAQORGSILYQCDYSTGSCPEI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGLCFLFGSNLRQOPQK 136  
Qy 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFPRMKWVSTVMEQLKSKTILFSLMQYSEEF 180  
Db 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFPRMKWVSTVMEQLKSKTILFSLMQYSEEF 196  
Qy 181 RIHFTKFEFONNPNRSLIPIITQLGRTHTATGLRKVRELFINITNGARKNAFKILFL 240  
Db 197 RIHFTKFEFONNPNRSLIPIITQLGRTHTATGLRKVRELFINITNGARKNAFKILVVI 256  
Qy 241 TDGEKGDPLAGYEDVPELDRGVIRVVGWDAFSEKSRQELNVTASKPPDRHVFQIN 300  
Db 257 TDGEKGDPLAGYEDVPELDRGVIRVVGWDAFSEKSRQELNVTASKPPDRHVFQIN 316  
Qy 301 NFEALKTIONQRLREKIFAIRGTQTGSSSSFEHEMSQGFSAAITNSGPLLSTVGSYDWAG 360  
Db 317 NFEALKTIONQRLREKIFAIRGTQTGSSSSFEHEMSQGFSAAITNSGPLLSTVGSYDWAG 376  
Qy 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYQHIGLVAMFR 436  
Qy 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSCPL 480  
Db 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSCPL 496  
Qy 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 540  
Db 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAYVLF 556  
Qy 541 HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSORIASGKLSPLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
Qy 601 PVLRVKAIMBFPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 660  
Db 617 PVLRVKAIMBFPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736  
Qy 721 SLVGTPLSAGFNLRPVLAEADAQRLFTALFPPEKNCNDNICQDDLSITFFSMDLCLVVG 780  
Db 737 SLVGTPLSAGFNLRPVLAEADAQRLFTALFPPEKNCNDNICQDDLSITFFSMDLCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900

Db 857 SGALKSTSCSINHPIFFPENSESVTFNITFDVDSKASLGKLLKLLKANVTSENNNPRTNKTEF 916  
QY 901 QLELPVKYAVVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSLPLSLVFLVPV 960  
Db 917 QLELPVKYAVVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSLPLSLVFLVPV 976  
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1020  
Db 977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1036  
QY 1021 FPGIQEENATLKNLSFDWYIKTSHNHLIIIVSTAEIILPNDVSFTLLPGQGAFAVRSQTET 1080  
Db 1037 FPGIQEENATLKNLSFDWYIKTSHNHLIIIVSTAEIILPNDVSFTLLPGQGAFAVRSQTET 1096  
QY 1081 KVEPFEVNPPLIIVGSSVGGLLLIALLIATAALYKLGFFKQYKQKMMSEGGPPGABPQ 1137  
Db 1097 KVEPFEVNPPLIIVGSSVGGLLLIALLIATAALYKLGFFKQYKQKMMSEGGPPGABPQ 1153

RESULT 11

US-08-476-062A-43  
; Sequence 43, Application US/08476062A  
; Patent No. 5877275  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY  
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476.062A  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/216,081  
; FILING DATE: 21-MAR-1994  
; APPLICATION NUMBER: 07/637,830  
; FILING DATE: 04-JAN-1991  
; APPLICATION NUMBER: 07/539,842  
; FILING DATE: 18-JUN-1990  
; APPLICATION NUMBER: 07/212,573  
; FILING DATE: 28-JUN-1988

; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00786/068003  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1152 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-476-062A-43

Query Match 98.9%; Score 5821.5; DB 2; Length 1152;  
Best Local Similarity 98.9%; Pred.No. 0;  
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGQSVOVLOGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCBPI 60  
Db 17 FNLDTENAMTFOENARGFGQSVOVLOGSRVVVGAPOEIVAAQNRGSLYQCDYSTGSCBPI 76  
QY 61 RLQVPEAVNMSGLSLAATTSPQLLACGPTVHTCSENTYVVKGLCLFLGSLNRQQPQK 120  
Db 77 RLQVPEAVNMSGLSLAATTSPQLLACGPTVHTCSENTYVVKGLCLFLGSLNRQQPQK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEWVSTVMEQLKSKTILFSLMQYSEBF 180  
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTILFSLMQYSEBF 196  
QY 181 RIHFTFKFQNNPNSRLIKPTTOLLGRTHATGLRKVVRELFINITNGARKNAFKILFLL 240  
Db 197 RIHFTFKFQNNPNSRLIKPTTOLLGRTHATGLRKVVRELFINITNGARKNAFKILVVI 256  
QY 241 TDGEKFGDPLGYEDVIPELDRGVIRYVVGDAFRSEKSRQELMTVASKPPRDHVFQIN 300  
Db 257 TDGEKFGDPLGYEDVIPADREGVIRYVVGDAFRSEKSRQELMTIASKPPRDHVFQIN 316  
QY 301 NFEALKTIQNLREKIFAIEGTOTGSSSSSFEHMSOEGFSAAITSGNPLISTVGSYDWAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTOTGSSSSSFEHMSOEGFSAAITSGNPLISTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLAGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLAGAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCP 480  
Db 437 QNTGMWESNANVKGTQIGAYGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSVCP 496  
QY 481 PRGQARWQCDVILYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGALGAEEDNRGAAYLF 540  
Db 497 PRG-RARWQCDVILYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGALGAEEDNRGAAYLF 555  
QY 541 HGTSGSGISPSHSORISAGSKLSPLQYFGQSLSGGDLTMDGLVDTLVGAQGHVLLRSQ 600  
Db 556 HGTSGSGISPSHSORISAGSKLSPLQYFGQSLSGGDLTMDGLVDTLVGAQGHVLLRSQ 615  
QY 601 PVLRVKAIIMEFNPREVARNVFECDNDQVYVKGAEGRVYCLHVQKSTRDLREGQIQSVVT 660  
Db 616 PVLRVKAIIMEFNPREVARNVFECDNDQVYVKGAEGRVYCLHVQKSTRDLREGQIQSVVT 675  
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720  
Db 676 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 735  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKKGNDNICODDLISITFSFMSLDCLVVG 780  
Db 736 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPPEKKGNDNICODDLISITFSFMSLDCLVVG 795  
QY 781 GPREFNVTIVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASSTEV 840  
Db 796 GPREFNVTIVRNDGEDSYRTQVTFPPDLISYRKVSTLQNRQSRWRLACESASSTEV 855  
QY 841 SGALKSTSCSINHPIFFPENSESVTFNITFDVDSKASLGKLLKLLKANVTSENNNPRTNKTEF 900  
Db 856 SGALKSTSCSINHPIFFPENSESVTFNITFDVDSKASLGKLLKLLKANVTSENNNPRTNKTEF 915  
QY 901 QLELPVKYAVVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSLPLSLVFLVPV 960  
Db 916 QLELPVKYAVVMVTSYKYNFTASENTSRVMQHOYQVSNLQORSLPLSLVFLVPV 975  
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1020  
Db 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP 1035  
QY 1021 FPGIQEENATLKNLSFDWYIKTSHNHLIIIVSTAEIILPNDVSFTLLPGQGAFAVRSQTET 1080  
Db 1036 FPGIQEENATLKNLSFDWYIKTSHNHLIIIVSTAEIILPNDVSFTLLPGQGAFAVRSQTET 1095  
QY 1081 KVEPFEVNPPLIIVGSSVGGLLLIALLIATAALYKLGFFKQYKQKMMSEGGPPGABPQ 1137

Db 1096 KVEPFEVNPPLVIGSSVGGLLALLIATAALYKLGFFKQYKDMMSGGPPGAEPQ 1152  
RESULT 12  
PCT-US96-01314-43  
; Sequence 43, Application PC/TUS9601314  
; GENERAL INFORMATION:  
; APPLICANT: M. Amin Arnaout  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN  
; TITLE OF INVENTION: ANTAGONISTS  
; NUMBER OF SEQUENCES: 78  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US96/01314  
; FILING DATE: 30-JAN-96  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/380,167  
; FILING DATE: 30-JAN-95  
; ATTORNEY/AGENT INFORMATION:  
; NAME: John W. Freeman  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00786/267001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1152  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; PCT-US96-01314-43  
Query Match 98.9%; Score 5821.5; DB 5; Length 1152;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;  
Qy 1 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQETVAANQSGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQETVAANQSGSLYQCDYSTGSCPEI 76  
Qy 61 RLQVPVEAVNMSLGLSLAAATSPQLLACGPTVHTQTCSENTYVVKGLCFLFGSNLRQPOQK 120  
Db 77 RLQVPVEAVNMSLGLSLAAATSPQLLACGPTVHTQTCSENTYVVKGLCFLFGSNLRQPOQK 136  
Qy 121 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEWSTVMEQLKSKTLFSLMOYSEEF 180  
Db 137 FPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLFSLMOYSEEF 196  
Qy 181 RIHETFEFQNNPNSRLIKETIQLGRTHTATGLRKVVRELFINITNGARKNAKILFLL 240  
Db 197 RIHETFEFQNNPNSRLIKETIQLGRTHTATGLRKVVRELFINITNGARKNAKILFLL 256  
Qy 241 TDGEKFGDPLGYEDVIPDLDEGVIRYVIGVDAPFRSEKSRQELNTVASKPRDHVFOIN 300  
Db 257 TDGEKFGDPLGYEDVIPDLDEGVIRYVIGVDAPFRSEKSRQELNTVASKPRDHVFOIN 316  
Qy 301 NFEALKTIQNLREKIPAEIQTGTGSSSSFEHMSQEGFSAATSNGLLSTVGSYDWAG 360  
Db 317 NFEALKTIQNLREKIPAEIQTGTGSSSSFEHMSQEGFSAATSNGLLSTVGSYDWAG 376

Qy 361 GVPLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 420  
Db 377 GVPLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVOSLVLGAPRYOHIGLVAMFR 436  
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYTYEQTRGGVSVCP 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYTYEQTRGGVSVCP 496  
Qy 481 PRQARWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPEGEDNRGAVLYP 540  
Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPEGEDNRGAVLYP 555  
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
Db 556 HGTSGSGISPSHSQRIAGSKLSPLOYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 615  
Qy 601 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660  
Db 616 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 675  
Qy 661 YDLALDSGRPHSRAVFNETKNSRTRQTVLGLTOTCETLKLQLPNCIEDPVSPIVLIN 720  
Db 676 YDLALDSGRPHSRAVFNETKNSRTRQTVLGLTOTCETLKLQLPNCIEDPVSPIVLIN 735  
Qy 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKMGNDNICQDDLSITTFMSLDCLVVG 780  
Db 736 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKMGNDNICQDDLSITTFMSLDCLVVG 795  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
Db 796 GPRESNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 855  
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Db 856 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915  
Qy 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSLSISLVLFPV 960  
Db 916 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMOHOYQVSNLQORSLSISLVLFPV 975  
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020  
Db 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1035  
Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080  
Db 1036 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTET 1095  
Qy 1081 KVEPFEVNPPLVIGSSVGGLLALLIATAALYKLGFFKQYKDMMSGGPPGAEPQ 1137  
Db 1096 KVEPFEVNPPLVIGSSVGGLLALLIATAALYKLGFFKQYKDMMSGGPPGAEPQ 1152  
RESULT 13  
5424399-2  
; Patent No. 5424399  
; APPLICANT: ARNAOUT, M. AMIN  
; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/78,871  
; FILING DATE: 16-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 539,842  
; FILING DATE: 18-JUN-1990  
; APPLICATION NUMBER: 212,573  
; FILING DATE: 28-JUN-1988  
; SEQ ID NO:2:  
; LENGTH: 1152  
; 5424399-2  
Query Match 98.9%; Score 5821.5; DB 6; Length 1152;



Best Local Similarity 98.9%; Pred. No. 0;		Matches 1125; Conservative 7; Mismatches 4; Indels 1; Gaps 1;	
QY	1	FNLDTENAMTFQENARGFQGSVVQIQGSRVVVVGAPQEIIVAAANQKSLYQCDYSTGSCBPI	60
Db	17	FNLDTENAMTFQENARGFQGSVVQIQGSRVVVVGAPQEIIVAAANQKSLYQCDYSTGSCBPI	76
QY	61	RLQVPEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLQOQOK	120
Db	77	RLQVPEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLQOQOK	136
QY	121	FPEARLGCPEQSDIAFLIDGSGSIIPHDFFRMKEMVSTVMBQLKKSKTLFSLMOYSEEF	180
Db	137	FPEARLGCPEQSDIAFLIDGSGSIIPHDFFRMKEMVSTVMBQLKKSKTLFSLMOYSEEF	196
QY	181	RIHFTFKPEQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAKFLILL	240
Db	197	RIHFTFKPEQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAKFLILVI	256
QY	241	TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGDAFRSEKSRQBLNTVASKPPRDHVQIN	300
Db	257	TDGEKFGDPLGYEDVPELDREGVIRYVIGVDGDAFRSEKSRQBLNTIASKPPRDHVQVN	316
QY	301	NFEALKTIONQREKIFAIEGTQTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDMAG	360
Db	317	NFEALKTIONQREKIFAIEGTQTGSSSFHEMSQEGFSAAITNSGPLLSTVGSYDMAG	376
QY	361	GVFLYTSKESKSTFINNRVDSMDNDAYLGAAATILNRVQSLVILGAPRYQHIGLVAMPR	420
Db	377	GVFLYTSKESKSTFINNRVDSMDNDAYLGAAATILNRVQSLVILGAPRYQHIGLVAMPR	436
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCLP	480
Db	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCLP	496
QY	481	PRGQARWQCDVAVLGEQGPWGRFGAALTIVLGDVNGDKLTDAIAPGAEEDNRGAVILF	540
Db	497	PRG-RARWQCDVAVLGEQGPWGRFGAALTIVLGDVNGDKLTDAIAPGAEEDNRGAVILF	555
QY	541	HGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGVLDTVGAQGHVILLRSQ	600
Db	556	HGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGVLDTVGAQGHVILLRSQ	615
QY	601	PVLRVKAIEMFNPREVARNVFNCNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
Db	616	PVLRVKAIEMFNPREVARNVFNCNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	675
QY	661	YDLALDSGRPHSRVAFNETKNSRRQTQVLGTQTCETLKLQLPNCIEDPVSPVILRLNF	720
Db	676	YDLALDSGRPHSRVAFNETKNSRRQTQVLGTQTCETLKLQLPNCIEDPVSPVILRLNF	735
QY	721	SLVGTPLSAFGLNLRPLVAEDAQRLETFALPPFEKNCNDNICODDLSITFSFMSLCLVVG	780
Db	736	SLVGTPLSAFGLNLRPLVAEDAQRLETFALPPFEKNCNDNICODDLSITFSFMSLCLVVG	795
QY	781	GPREENVTVVRNDGEDSVRTQVTFPPFLDLRYKVTSLQNSQRSWRLACESASSTEV	840
Db	796	GPREENVTVVRNDGEDSVRTQVTFPPFLDLRYKVTSLQNSQRSWRLACESASSTEV	855
QY	841	SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNPRNKTEF	900
Db	856	SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNPRNKTEF	915
QY	901	QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVWQHOYQVSNLQORSPLSLVFLVPV	960
Db	916	QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVWQHOYQVSNLQORSPLSLVFLVPV	975
QY	961	RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP	1020
Db	976	RLNQTVIWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDDIP	1035
QY	1021	PFGIQEFENATLKGNSLSPDWYIKTSHNHLIIYSTAAILFNDVSFTLLPQCGAFVRSQTET	1080

Db	1036	PFGIQEFENATLKGNSLSPDWYIKTSHNHLIIYSTAAILFNDVSFTLLPQCGAFVRSQTET	1095
QY	1081	KVEPEVNPPLIYVSSVGGILLALITAAALYKLGFKROVKDMMSEGGPGABPQ	1137
Db	1096	KVEPEVNPPLIYVSSVGGILLALITAAALYKLGFKROVKDMMSEGGPGABPQ	1152

RESULT 14

US-08-476-062A-44

Sequence 44, Application US/08476062A

Patent No. 5877275

GENERAL INFORMATION:

APPLICANT: Arnaout, M. Amin

TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY

TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESS: Fish & Richardson P.C.

STREET: 225 Franklin Street

CITY: Boston

STATE: MA

COUNTRY: US

ZIP: 02110-2804

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: Windows95

SOFTWARE: FastSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/476,062A

FILING DATE: 07-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/216,081

FILING DATE: 21-MAR-1994

APPLICATION NUMBER: 07/637,830

FILING DATE: 04-JAN-1991

APPLICATION NUMBER: 07/539,842

FILING DATE: 18-JUN-1990

APPLICATION NUMBER: 07/212,573

FILING DATE: 28-JUN-1988

ATTORNEY/AGENT INFORMATION:

NAME: Freeman, John W.

REGISTRATION NUMBER: 29,066

REFERENCE/DOCKET NUMBER: 00786/068003

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906

TELEX: 200154

INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:

LENGTH: 1163 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-476-062A-44

Query Match 58.8%; Score 3459; DB 2; Length 1163;

Best Local Similarity 60.9%; Pred. No. 1.2e-289;

Matches 687; Conservative 142; Mismatches 294; Indels 6; Gaps 4;

QY	1	FNLDTENAMTFQENARGFQGSVVQIQGSRVVVVGAPQEIIVAAANQKSLYQCDYSTGSCBPI	60
Db	20	FNLDTENAMTFQENARGFQGSVVQIQGSRVVVVGAPQEIIVAAANQKSLYQCDYSTGSCBPI	79
QY	61	RLQVPEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLQOQOK	120
Db	80	GLQVPEAVNMSLGLSLAATTPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLQOQOK	137
QY	121	FPEARLGCPEQSDIAFLIDGSGSIIPHDFFRMKEMVSTVMBQLKKSKTLFSLMOYSEEF	180
Db	138	LPVSQECPEQSDIAFLIDGSGSIIPHDFFRMKEMVSTVMBQLKKSKTLFSLMOYSEEF	197
QY	181	RIHFTFKPEQNNPRSLIKPTTOLLGRTHATGLRKVVRELFNITNGARKNAKFLILL	240





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds  
(without alignments)  
4014.622 Million cell updates/sec

Title: US-09-902-481b-5  
Perfect score: 5876  
Sequence: 1 FNLDTENAMTFQENARGFQ.....FKRQYKDMSEGGPGGAEFQ 1137  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	1 RWHU1B	cell surface glyco
2	4470	76.1	1153	2 S00551	leukocyte surface
3	3479	59.2	1163	1 RWHU1C	cell surface glyco
4	1533.5	26.4	1170	2 S03308	cell surface glyco
5	1538.5	26.2	1163	2 I56126	lymphocyte fuction
6	1153	19.6	1179	2 A53213	integrin alpha-E c
7	1103.5	18.8	1151	2 A45226	integrin alpha-1 c
8	1085	18.5	1170	2 I45914	integrin alpha 2 s
9	1071	18.2	1178	2 S44142	VLA-2 protein homo
10	1068	18.2	1181	2 A33998	integrin alpha-2 c
11	1062	18.1	1180	2 A35854	integrin alpha-1 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	635	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-9 c
15	614.5	10.5	1041	2 T31437	integrin alpha cha
16	579.5	9.9	1054	2 J07294	alpaP integrin -
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	VLA-3 alpha subuni
19	555.5	9.5	1053	2 S44250	integrin alpha-5 c
20	542.5	9.2	1034	2 A36108	integrin alpha-v c
21	535	9.1	1044	2 T10050	integrin alpha-v c
22	532	9.1	1049	2 A27079	integrin alpha-v c
23	532	9.1	1073	2 B36429	fibronectin recept
24	531.5	9.0	1072	2 A38457	integrin alpha-6 c
25	530.5	9.0	1051	2 A40021	integrin alpha-6 c
26	526	8.9	1048	2 A27421	integrin alpha-3 c
27	525.5	8.9	1091	2 A41543	integrin alpha-5 c
28	517	8.8	1044	2 S16516	integrin alpha-6 c
29	496.5	8.4	1146	2 S40311	integrin - fruit f

## ALIGNMENTS

## RESULT 1

## RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Ma  
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004  
R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.  
J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CR  
B.

A:Reference number: A31108; MUID:88315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:Cross-references: UNIPROT:P11215; GB:J03925; NID:gl87284; PIDN:AAA59544.1; PID:g307148  
A:Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor M  
A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499,501-965,'P',967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAA594  
A:Note: the authors translated the codon TAC for residue 1129 as Thr

A:Note: part of this sequence, including the amino end of the mature protein, was confir  
R:Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally reg  
A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:gl80018; PIDN:AAA58410.1; PID:g553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesi  
A:Reference number: A94193; MUID:88190151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor  
A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:Cross-references: GB:J04145; NID:gl89068; PIDN:AAA59903.1; PID:g386975

A;Note: part of this sequence was confirmed by protein sequencing  
R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.  
J. Immunol. 150, 480-490, 1993  
A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in  
n during evolution.  
A;Reference number: A46526; MUID:93123748; PMID:8419480  
A;Accession: A46526  
A;Status: not compared with conceptual translation  
A;Molecule type: DNA  
A;Residues: 1-499,501-1153 <FILE>  
A;Cross-references: GB:S52227; NID:G263047; PIDN:AA824821.1; PID:G263049  
A;Note: the last three bases of intron 13, CAG, are included in some but not all mature  
A;Note: sequence extracted from NCBI backbone (NCBIP:121963)  
R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.  
Biochim. Biophys. Acta 874, 368-371, 1986  
A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp  
A;Reference number: A90664; MUID:87076671; PMID:3539202  
A;Accession: A26091  
A;Molecule type: protein  
A;Residues: 17-31 <PIE>  
A;Experimental source: granulocytes  
R;Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G.  
Blood 79, 865-870, 1992  
A;Title: Characterization of the myeloid-specific CD11b promoter.  
A;Reference number: I52567; MUID:12144986; PMID:1346576  
A;Accession: I52567  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-9 <RES>  
A;Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:G553219  
C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
C;Genetics:  
A;Gene: GDB:ITGAM; CR3A  
A;Cross-references: GDB:120599; OMIM:120980  
A;Map position: 16p11.2-16p11.2  
A;Note: promoter contains a GATA motif and two Spl consensus binding sites  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>  
F;17-1108/Domain: extracellular #status predicted <EXT>  
F;148-318/Domain: von Willebrand factor type A repeat homology <VMA2>  
F;465-473/Region: calcium/magnesium binding #status Predicted  
F;530-538/Region: calcium/magnesium binding #status Predicted  
F;533-601/Region: calcium/magnesium binding #status Predicted  
F;1109-1134/Domain: transmembrane #status predicted <TM>  
F;1135-1153/Domain: intracellular #status predicted <INT>  
F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.8%; Score 5862; DB 1; Length 1153;  
Best Local Similarity 99.3%; Pred No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGSVVQLQSGRVVVGAPQEIIVAAANQSGSLYQCDYSTGCEPI 60  
Db 17 FNLDTENAMTFQENARGFGSVVQLQSGRVVVGAPQEIIVAAANQSGSLYQCDYSTGCEPI 76

Qy 61 RLQVPVEAVNMISGLSLAATTPPOLLACGPTVHTQTSNTYVVKGLCFLFGSNLRQOPQK 120  
Db 77 RLQVPVEAVNMISGLSLAATTPPOLLACGPTVHTQTSNTYVVKGLCFLFGSNLRQOPQK 136

Qy 121 FPEALRGCPQSDIAFLVDGSGSIIIPHDPRAKEFTSTVMEQLKSKTLFSLMQYSEEF 180  
Db 137 FPEALRGCPQSDIAFLVDGSGSIIIPHDPRAKEFTSTVMEQLKSKTLFSLMQYSEEF 196

Qy 181 RIHFTTFEQNNPRSLIKPITQLLGRTHATGIRKVVRELFININGARQNAFKILILI 240  
Db 197 RIHFTTFEQNNPRSLIKPITQLLGRTHATGIRKVVRELFININGARQNAFKILIVI 256

Qy 241 TDGEKFGPLGYEDVIPADREGVIRYVIGVGDAFRSEKSKQELNTVASKPPRDHVFOIN 300  
Db 257 TDGEKFGPLGYEDVIPADREGVIRYVIGVGDAFRSEKSKQELNTIASKPPRDHVFOIN 316

Qy 301 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPIILLSTVGSVDWAG 360  
Db 317 NFEALKTIONQLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPIILLSTVGSVDWAG 376

Qy 361 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAILLRNRVQSLVGLGAPYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINNTRVDSMDNAYLGAAAILLRNRVQSLVGLGAPYQHIGLVAMFR 436

Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSDVDNSNGSTDVLIGAPHYYEYTRGGQVSVCP 480  
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSDVDNSNGSTDVLIGAPHYYEYTRGGQVSVCP 496

Qy 481 PRQQRARWCDAVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAIGAPBEDNRGAVLYF 540  
Db 497 PRQQRARWCDAVLYGEOQPMGRFGAALTIVLDVNGDKLTDVAIGAPBEDNRGAVLYF 556

Qy 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616

Qy 601 PVLRVKAIMFENPREVARNVFECNDQVWKGKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660  
Db 617 PVLRVKAIMFENPREVARNVFECNDQVWKGKEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676

Qy 661 YDLALDSGRPHSAVFNETKNSTRRTQVLGLTQTCTETLKLQLPNCIEDPVSIVLRLNF 720  
Db 677 YDLALDSGRPHSAVFNETKNSTRRTQVLGLTQTCTETLKLQLPNCIEDPVSIVLRLNF 736

Qy 721 SLVGTPLSAGNIRPVLAEADAORLFTALPPEKNCNDNICODDLSITTFMSLDCLVVG 780  
Db 737 SLVGTPLSAGNIRPVLAEADAORLFTALPPEKNCNDNICODDLSITTFMSLDCLVVG 796

Qy 781 GPREENVTVRNDGEDSVYRTQVTFPPDLDSYRKVSTLQNRQSRSWRLACESASSTEV 840  
Db 797 GPREENVTVRNDGEDSVYRTQVTFPPDLDSYRKVSTLQNRQSRSWRLACESASSTEV 856

Qy 841 SGALKSTSCSINHPIPPENSEVTNFTFDVDSKASLGKLLKANVTSENNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIPPENSEVTNFTFDVDSKASLGKLLKANVTSENNMPTNKTEF 916

Qy 901 QLELPKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVSNLQORSILPISLVFLVPV 960  
Db 917 QLELPKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVSNLQORSILPISLVFLVPV 976

Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP 1036

Qy 1021 FPGIQEENATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOTET 1080  
Db 1037 FPGIQEENATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSOTET 1096

Qy 1081 KVBPFEVNPPLPLTVGSSVGLLLALITAAALYKGFQKQYKDMSEGPPGAEPO 1137  
Db 1097 KVBPFEVNPPLPLTVGSSVGLLLALITAAALYKGFQKQYKDMSEGPPGAEPO 1153

RESULT 2

S00551

Leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N;Alternate names: complement-3 receptor alpha chain

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C;Accession: S00551; I59078

R;PyTela, R.

EMBO J. 7, 1371-1378, 1988

A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the

A;Reference number: S00551; MUID:88312584; PMID:3044779

A;Accession: S00551

A;Molecule type: DNA

A;Residues: 1-1153 <PVT>

A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:G52982; PIDN:CAA30479.1; PID:G5298

A;Note: the authors translated the codon CAC for residue 569 as Glu

R;Sastre, L.; Roman, J.M.; Teplov, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986  
A;Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep  
A;Reference number: 159078; MUID:86287312; PMID:2942940  
A;Accession: 159078  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 11-44 <RES>  
A;Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193  
C;Genetics:  
A;Gene: Mac-1  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C;Keywords: cell adhesion; glycoprotein; transmembrane protein  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental  
F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>  
F;1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.1%; Score 4470; DB 2; Length 1153;  
Best Local Similarity 73.9%; Pred. No. 7-9e+303;  
Matches 841; Conservative 146; Mismatches 149; Indels 2; Gaps 2;

QY 1 ENLDTENAMTFQENARFGQSVVOLQGRVVRVGAPOEIVAAVNRGLSYQCDYSTGSCPEI 60  
Db 17 ENLDTENAMTFQENARFGQSVVOLQGRVVRVGAPOEIVAAVNRGLSYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQCSTENTYKGLCFLEGSNLRQOPQK 120  
Db 77 PLQVPVEAVNMSLGLSLAVSTVPQQLLACGPTVHQCSTENTYKGLCFLEGSNLRQOPQK 136

QY 121 FPEALRGCPQSDIAFLVDGSGSIIPHDPRAKEFISTVMEQKKSTLFSLMQYSEEF 180  
Db 137 FPEALRGCPQSDIAFLVDGSGSIIPHDPRAKEFISTVMEQKKSTLFSLMQYSEEF 196

QY 181 RHFTFKFQNNPNSLKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
Db 197 RHFTFKFQNNPNSLKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRORLNTVASKPRDHFVQIN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRORLNTVASKPRDHFVQIN 316

QY 301 NFEALKTIQNLREKIFAIEGTQTSSTSSFEHMSQEGFSAITSNGLPSTVGSVDWAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTQTSSTSSFEHMSQEGFSAITSNGLPSTVGSVDWAG 376

QY 361 GVFLYTSKEKSTFNNTRVDSMDNAYLGAAAILRNVRQSLVIGAPRYQHIGLVAMPR 420  
Db 377 GAFLYTSKDKVTFINTRVDSMDNAYLGAAAILRNVRQSLVIGAPRYQHIGLVAMPR 436

QY 421 QNTGMESNANVKGQTIGAYFGASLCSVDVDSNGSTDLVIGAPHYETRGQSVVCP 480  
Db 437 ENFGTWEHTSIKSGQISYFGASLCSVDVDSNGSTDLVIGAPHYETRGQSVVCP 496

QY 481 PRGQARWQCDVAVLGEQGPWRGAALTVLGDVNGDKLTDVAICAGEEDNRGAVILF 540  
Db 497 PRGQARWQCDVAVLGEQGPWRGAALTVLGDVNGDKLTDVAICAGEEDNRGAVILF 555

QY 541 HCTSGSGISPHSRIAGSKLSPRIQYFGQSLGQGLTMDGLVDLTVCAQGHVLLRSQ 600  
Db 556 YGASIASLASHSHRIIGAHPFGLQYFGQSLGQGLTMDGLVDLTVCAQGHVLLRSQ 615

QY 601 PVLVRKALMEFNPVARNVFCNDQVVKGEKAGEVRVCLVHVKSTRLRGQOSVVT 660  
Db 616 PVLVRKALMEFNPVARNVFCNDQVVKGEKAGEVRVCLVHVKSTRLRGQOSVVT 675

QY 661 YDLALDSGRPHSVAFNETKSTRRTQVGLGTQTCETLKLPLNCIEDPVPVILRLNF 720  
Db 676 YDLALDSGRPHSVAFNETKSTRRTQVGLGTQTCETLKLPLNCIEDPVPVILRLNF 735

QY 721 SLVGTPLSAGNLRVLAEDAQRLETFALPPFEKNGCNNDICODDLSITFSFMSLCLVVG 780  
Db 736 TLVGEPLRSFNGLRVLAEDAQRLETFALPPFEKNGCNNDICODDLSITFSFMSLCLVVG 795

RESULT 3  
RWHLIC

N;Alternate names: leukocyte adhesion precursor - human

C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text\_change 09-Jul-2004

C;Accession: A35543; A35543; S00864

R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 12750-12751, 1990

A;Reference number: A36584

A;Contents: erratum

A;Accession: A36584

A;Molecule type: DNA

A;Residues: 1-1163 &lt;COR&gt;

A;Cross-references: UNIPROT:P20702

A;Note: this revision to the sequence from reference A35543 includes the carboxyl end

R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.

J. Biol. Chem. 265, 2782-2788, 1990

A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.

A;Reference number: A35543; MUID:90153906; PMID:2303426

A;Accession: A35543

A;Molecule type: DNA

A;Residues: 1-834 &lt;CO2&gt;

A;Note: this sequence has been revised in reference A36584

R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.

EMBO J. 6, 4023-4028, 1987

A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte

A;Reference number: S00864; MUID:88166645; PMID:3327687

A;Accession: S00864

A;Molecule type: mRNA

A;Residues: 1-755, 'L', 757-1163 &lt;CO3&gt;

A;Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830

A;Note: part of this sequence was confirmed by protein sequencing

C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my

C;Genetics:

A;Gene: GDB:ITGAX; CD11C

A;Cross-references: GDB:119758; OMIM:151510

A;Map position: 16p11.2-16p11.2

C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom

C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;

F;1-19/Domain: signal sequence #status predicted &lt;SIG&gt;

F;20-1163/Product: cell surface glycoprotein CD11c #status predicted &lt;MAT&gt;

F;20-1107/Domain: extracellular #status predicted &lt;EXT&gt;

F;1108-1133/Domain: von Willebrand factor type A repeat homology &lt;VWA4&gt;

F;1134-1163/Domain: transmembrane #status predicted &lt;TM&gt;

F;61.89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 59.2%; Score 3479; DB 1; Length 1163;  
Best Local Similarity 61.1%; Pred. No. 9.3e-234;  
Matches 690; Conservative 141; Mismatches 292; Indels 6; Gaps 4;

QY 1 ENLDTENAMTFQENARGFGSGVVLQGSRRVVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60  
DB 20 ENLDTBELTAFRVDSAGFGSVVQYANVWVVGAPQKITANQGTGGLYQCYGTGACEPI 79  
QY 61 RLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVVKGLCFPLFGSNLRQOPOK 120  
DB 80 GLQVPPEAVNMSLGLSLASTTSPQLLACGPTVHHEGRNMYLTGLCFLLGPT--QLTOR 137  
QY 121 FPAALRGCPQEDSDIAFLVDGSGSIIIPDFRRAKEFTTMEQLKSKTILFSLMOYSEEF 180  
DB 138 LPVSRQECPRQEQDVLIDGSGSISRNFAFMNFVRAVISQFQRPSTQFSLMQFSNKF 197  
QY 181 RIHFTKFEQNNPNRSLIKPITOLLGRTHATGIRKVVRELFINITGARKNAFKILILI 240  
DB 198 QTHFTFEFRRTSNPLSLLASVHQLQGFTYTATAIQNVHRLPHASVGARDATKILIVI 257  
QY 241 TDGEKFGDPLGYEDVIEPADREGVIRYVIGVGDAPRSEKSRQELNITVASKPPRDHVFQIN 300  
DB 258 TDGKKEGSLDYKDVIPMAAAGIIRYAIGVGLAFQNRKNGKWLNDIASKPSQEHIEKVE 317  
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSPHEHMSQEGFSAATITNSGFLLSYVGSYDWAG 360  
DB 318 DFDALKDIONLREKIFAIEGTQTSSTSSFELEMAQEGFSAVFPDGPVILGAVGSFTWSG 377  
QY 361 GVFLYTSKESKSTFINMTFRVDSMDMDAYLGAIAAILNRVQSLVGLGAPRYQHIGLVAMFR 420  
DB 378 GAFLYPPNMSPTFINMSQENVMDSDVLYGYSYTELALWKGVQSLVGLGAPRYQHTGKAVIFT 437  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYQEYTRGGQVSCVPL 480  
DB 438 QVSRQWRKAEVTTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYQEYTRGGQVSCVPL 497  
QY 481 PRGORARWCDVAVLYGQGPQWGRGALTVLGVNDGDKLTDVAIGAPGEDNKGAVLYF 540  
DB 498 PRGWR--RWCDAVLYGEGHGWGFGAALTVLGVNDGDKLTDVVGAPGEENKGAVLYF 556  
QY 541 HGTSGSGISPSHSORIASGLSPRLQFQSGLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
DB 557 HGVLPGISPSHSORIASGLSRLQFQSGLSGGQDLTQGLVDLTVGARGQVLLRTR 616  
QY 601 PVLRVKAIMFNPREVARNFECDNVVKGKEAGEVRVCLHVQKSTRDRLREGQIQSVVT 660  
DB 617 PVLWVGVSQMQIPAEIPRSAFECEQVVSQETLVQSNICLYIDKRSKNLLGSRDLQSSVT 676  
QY 661 YDLALDSGRPHSRVAFVNETKNSRTOVLGHTQTCETLKLQLPNCIEDPVSPIVLRLNF 720  
DB 677 LDALDPGLRSPRATFQETKNSRVRVVLGLKAHCENFNLLPSCVEDSVTPITLRLNF 736  
QY 721 SLVGTPLSAFGLNPLVLAEDAQRLFTALFPPEKXNGNDNIQDDLSITFSFMSDCLVVG 780  
DB 737 TLVGKPLLAFLNPLMLAADAQRYFTASLPPEKNGADHIQDNLGISFSFPLGKSLVVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRORSORSLACESASSTEV 840  
DB 797 SNLEINAEVWVNDGEDSYGTTITFSHPAGLSYRYVAEGQKQGLRSLHLTCDSPVGV-- 854  
QY 841 SGALKSTSCSNHPTFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPNTRNKTKEF 900  
DB 855 SQGTWSTSCRNHLIFRGAQITFLATFDVSPKAVLGDRLLLLTANVSSENTPRTSKTTF 914  
QY 901 QLELPKVAVYVMTVSHGVSTKYLNFITAS--ENTSRVMOHQYVSNLQORSILPISILVFLVP 959  
DB 915 QLELPKVAVYVTVSSHEQFTKYNLFSSEEEKESHVAMHRYQVNNLQORDLPVSNFWVP 974  
QY 960 VRLNQTIVDRPQVTFSENLSCTHTKERLPHSHDFLAELKAPVNVGSIACVQRLQDI 1019  
DB 975 VELNQAEMVMDVEVSHPPNPSLRCSSEKIAPPASDPFLAHIQKPNVLDCSIACGLFRCDV 1034

QY 1020 PFGIQEENFATLKNLSFDWIKTSHNHLIIVSTAEILFNDSVFTLLPFGQGFVRSQTE 1079  
DB 1035 PPSFVQGEELDFTLKGNLSFGWVRQIIQKKVSVVVAEITFTDTSVLSQLPQGEAFMRAQIT 1094  
QY 1080 TKVEPVEPNPUPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSE 1128  
DB 1095 TVLEKYKVHNPTPLIVGSSSTGGLLLLALITAVLYKVGVFFKQYKEMEE 1143

RESULT 4  
S03308  
cell surface glycoprotein CD11a precursor - human  
N;Alternate names: leukocyte adhesion glycoprotein LFA-1 alpha chain; leukocyte function  
C;Species: Homo sapiens (man)  
C;Date: 28-Feb-1990 #sequence revision 28-Feb-1990 #text\_change 09-Jul-2004  
C;Accession: S03308; A47458; A47565; A48759; S36044  
J;Larson, R.S.; Corbi, A.L.; Berman, L.; Springer, T.  
J;Cell Biol. 108, 703-712, 1989  
A;Title: Primary structure of the leukocyte function-associated molecule-1 alpha subunit  
A;Reference number: S03308; MUID:89139587; PMID:2537322  
A;Accession: S03308  
A;Molecule type: mRNA  
A;Residues: 1-1170 <LAR>  
A;Cross-references: UNIPROT:P20701; UNIPROT:Q9UBC8; EMBL:Y00796; NID:g31421; PIDN:CRA68  
A;Note: part of this sequence was confirmed by protein sequencing  
R;Cornwell, R.D.; Gollabon, K.A.; Hickstein, D.D.  
Proc. Natl. Acad. Sci. U.S.A. 90, 4221-4225, 1993  
A;Title: Description of the leukocyte function-associated antigen 1 (LFA-1 or CD11a) pro  
A;Reference number: A47458; MUID:93248261; PMID:8097887  
A;Accession: A47458  
A;Molecule type: DNA  
A;Residues: 1-20 <COR>  
A;Note: sequence extracted from NCBI backbone (NCBIN:130862, NCBI:P:130863)  
R;Shelley, C.S.; Farokhzad, O.C.; Arnaout, M.A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 5364-5368, 1993  
A;Title: Identification of cell-specific and developmentally regulated nuclear factors  
A;Reference number: A47565; MUID:93281759; PMID:8099450  
A;Accession: A47565  
A;Molecule type: DNA  
A;Residues: 1-20 <SHS>  
A;Cross-references: GB:M95609  
R;Nueda, A.; Lopez-Cabrera, M.; Vara, A.; Corbi, A.L.  
J. Biol. Chem. 269, 19305-19311, 1993  
A;Title: Characterization of the CD11a (alphaL, LFA-lalpha) integrin gene promoter.  
A;Reference number: A48759; MUID:93374910; PMID:8103515  
A;Accession: A48759  
A;Molecule type: DNA  
A;Residues: 1-20 <NUE>  
A;Cross-references: EMBL:Z22804; NID:g311405; PIDN:CAA80461.1; PID:g311406  
C;Genetics:  
A;Gene: GDB:ITGAL; CD11A  
A;Cross-references: GDB:119757; OMIM:153370  
A;Map position: 16p11.2-16p11.2  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C;Keywords: cell adhesion; cytoskeleton; glycoprotein; heterodimer; surface antigen; tr  
P;1-25/Domain: signal sequence #status predicted <SIG>  
P;26-1170/Product: leukocyte adhesion glycoprotein LFA-1 alpha chain #status predicted  
P;154-317/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 26.4%; Score 1553.5; DB 2; Length 1170;  
Best Local Similarity 34.7%; Pred. No. 1.5e-99;  
Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;

QY 1 ENLDTENAMTFQ--ENARGFGSVVVLQGSRRVVVVGAPQEIIVAAQNGSLYQCDYSTGSC 58  
DB 26 YNLDVGRARSFSPRAGHFYRVLQV--GNGVYVGAPE---GNSTGLYQCQSGTGHCL 81  
QY 59 PIRLQVPVEAVNMSLGLSLAATTPPQLLACGPTVHQTCSNTYVVKGLCFPLFGSNLR--- 115  
DB 82 PVTLLR--GSNYTSKYLGMTLATDPTDGSILACDPGLSRCTCDQNTYLSGLCYLFRQLQGP 140  
QY 116 --QPOKPFPEARLGCPEQSDSDIAFLVDGSGSIIIPHDFRAKSFISTVMSQLKSKTLFSLM 174



Db 141 LQGRFGQECIKG-----NVDLVFLFDGMSLQDFQKILDFMKDVMKKLSNTSYQFAAV 196  
 QY 175 YSEEFRIHFTKEFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAP 234  
 Db 197 QPSTSYKTEFDSDVVKWKDPALLKHVKHMLLINTGAINVATEVFEELGARPDAT 256  
 QY 235 KILLITDGEKFGDPLGVEDVIPADRGVIRYVIGVDAPRSEKSRQELNNTVASKPPRD 294  
 Db 257 KVLIIITDGE--ATDSGNIDAAD-----IIRYIIIGIKGHFQTKESQETLHKFASKPASE 309  
 QY 295 HVFOINNEALKTIQNLREKIPATEGTQGTSSSSFEHMSQEGFSAITNSGPIILSTVG 354  
 Db 310 FVKILDTFEKLDLTELQKKIYIIEGTSKODLTSFNMLSSGGSADLSRGAHVGAAG 369  
 QY 355 SYDWAGGVF-LYTSKEKSTFINMTRVDSMDMDAYLYGAAA-IILNRVQSLVLGAPRYQH 412  
 Db 370 AKDWAGGFLDLKADLQDDTFIGNPLTPEVRAGVLYGVTVLPSRQKSLASGAPRYQH 429  
 QY 413 IGLVAMFR--ONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEQT 470  
 Db 430 MGRVLLFOEPQGGHWSQVQTHGTQIGSYFGGELCGVDVDQDGETELLIGALPIFYGEQ 489  
 QY 471 RGGQVSCPLPRGORARWQCDVAV--LYGEOQOPMGRFGAALTVLGDVNGDKLTDVAIGAP 528  
 Db 490 RGRVFYI-----ORRLGFEVSELODGPYLPGRFGEAITALTIDNGDGLVDVAVGAP 544  
 QY 529 GEEDNRGAVLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTV 588  
 Db 545 LEE--QGANVIFNGRHG-GLSPQSPQRLEGTVLGSQWFGRSIHGVKDLGGDLADVAV 601  
 QY 589 GAQGHVLLLSQPVLRKALMEFNPREVARNVFECNDQV--KGKEAGEVRVCLHVQKSTR 647  
 Db 602 GAEQMTVLSSRPVDMVTILMSFPABIPVHEVECSYSTSNKMEGVNIITCFQI-KSLY 660  
 QY 648 DRLREGIQSVTVTDALDGRPHSRVAVENETKNSITRQTVLGLTQTCETLKLQLPNCI 707  
 Db 661 PQF-QGELVANLYTLQDGHTRRRGLFPGGRHELRNIAVT-TSMSCDTDFSFHFPCV 718  
 QY 708 EDVPSIVLRLNFSL---VGTPLS--AFGN-----LRPVLAEDAQRFTALFFPEKNCGN 757  
 Db 719 QDLISPINVSLNFSLWEEGTPRDQRAQGDIPILRPSLHSETWEI-----PFKNCGE 773  
 QY 758 DNTCQDDLSITFMSLDCLVGGPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVS 817  
 Db 774 DKCEANLRVSPARSRALRLTAFASLSVELSLNLEADAYVWQLDHPFGGLSRKVE 833  
 QY 818 TLQNRQSQRWRLACES--ASSTEVSALKSTSCSINHPIFPENSEVTFNITFDVDSKAS 875  
 Db 834 ML---KPHSQIPVSCPELPEERLLSRAL---SCNVSSPIFKAGSHVALQMENTLVNSS 887  
 QY 876 LGNKLKLLKANVTSENN---MPTNKTETFOLELPVKYAVVMVVTSHGVSTKYLNFTASEN 931  
 Db 888 WGDSDVELHANVTENNEDSDLLDENSAITTI--IPILYPINILIQOEDSTLVSFPTKGP 944  
 QY 932 TSRVMQHQVQV---SNLGRSLP-ISLVLVPLVRLNQTVIWRDPQVTFSENLSSTCHTK- 986  
 Db 945 KIHQVKHMYQVRQPSIHQHNITPLEAVVGVQPPSEGPITHQWSVOMBPV--PCHYED 1002  
 QY 987 -RLPSSHSD--FLAELRKAPVNCSTAVCQRQCDDIPFGIGQEEFNATLKNLSFDWYIK 1043  
 Db 1003 LERLPDAEPCLPGALFRCPVW-----FRQELVQVIGTLELVGEIE 1044  
 QY 1044 TSHNHLIVSTABIIFNDSVFTLLPQOGAFVRSQETKVEPEVFNPLPLIYVSSVGGLL 1103  
 Db 1045 AS-SMFSLCSLSISFNSSKHFLYGSNASL-AQVVMKVDDVYVEKQMLYLYVLSGIGLL 1102  
 QY 1104 LLALITAAIYKLGFFKQVQKQNMSEG-GPPGAP 1136  
 Db 1103 LLLLLIFVLYKGVFFKRNLERKWEAGRGVNGIP 1136

RESULT 5  
 156126

lymphocyte fuction-associated molecule-1-alpha - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
 C:Accession: 156126  
 R:Kaufmann, Y.; Tseng, E.; Springer, T.A.  
 J. Immunol. 147, 369-374, 1991  
 A:Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit  
 A:Reference number: 156126; MUID:91268576; PMID:2051027  
 A:Accession: 156126  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1163 <RES>  
 A:Cross-references: UNIPROT:P24063; GB:M60778; NID:g198785; PIDN:AAA39436.1; PID:g19878  
 C:Genetics:  
 C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
 F:151-315/Domain: von Willebrand factor type A repeat homology <WVAL>

Query Match 26.2%; Score 1538.5; DB 2; Length 1163;  
 Best Local Similarity 34.2%; Pred. No. 1.6e-98;  
 Matches 401; Conservative 214; Mismatches 458; Indels 99; Gaps 36;

QY 1 FNLDTENAMTFQENA-RGFGQSVVQLQGRVVVVGAPQBIIVAANQORGSILYQCDYSTGSCPE 59  
 Db 24 YNLDTRTQTFLAQAGRHFGYQLIEDG-VVVGAPGE---GDNTGGLYHCHRTSSEFCQP 79  
 QY 60 IRLQVPVEAVNMSLGLSLAATTPPQLLAGCTPVHQTCSNTYVYKGLCFLFGSNLRQOPQ 119  
 Db 80 VSLH-GSNHTSKYLGMTLATDAAKGSLACDGLSRTCDQNTYLSGLCYLPQSGLEGPML 138  
 QY 120 KFEALRGCPQESDSTAFIVDGGSGIIPHFRRAKEFISTVMEQLKKSKTLFSLMOYSEE 179  
 Db 139 QNRPAVQECMKGVLDVLFVFDGSQLDKRDKDFEKLFEPMKDMVRKLSNTSYQFAAVQFSD 198  
 QY 180 FRHFTFKPE-QNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAPKILI 238  
 Db 199 CRTEFTFLDVYKQKNKPDVLLGSVQPMFLNTFRAINVYVAVHVFKEESGARPDATKVLV 258  
 QY 239 LIITDGSKFGDPLGYEDVIPLEADREG-----VIRYVIGVDGAFRSEKSRQELNNTVASKP 291  
 Db 259 IITDG-----EASDKGNI SAADHITRIIIGIKGHFVSQKQKTLHIIFASEP 304  
 QY 292 PRDHVQIINNFEALKTIQNLREKIPATEGTQGTSSSSFEHMSQEGFSAITNSGPIILS 351  
 Db 305 VBEFVKILDTFEKLDLTDLQRIYAIERTNRQDLTSFNMLSSGGSADLSKGHAVVG 364  
 QY 352 TVGSDWAGGVF-LYTSKEKSTFINMTRVDSMDMDAYLYGAAA-IILNRVQSLVLGAPR 409  
 Db 365 AVGAKDWAGGFLDLREDLQCATFVGQEPDTSVDRGGYLGTYVAMWTSRSRPLLAAGAPR 424  
 QY 410 YHIGILVAMFR--ONTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHY 467  
 Db 425 YQHVGVLLFQAPAGGRWNQTKIEGTQIGSYFGGELCSVDLDQDGEALLIGALPLFF 484  
 QY 468 EOTRGQVSVCPPLPRGORARWQCDVAVLYGEOQOPMGRFGAALTVLGDVNGDKLTDVAIGA 527  
 Db 485 GEORGRVFTY---QRRQSLFEMVSELODGPYLPGRFGEAITALTIDNGDRLTDAVAGA 541  
 QY 528 POEEDNRGAVLFHGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLT 587  
 Db 542 PLEE--QGANVIFNGKPG-GLSPQSPQRLEGTVLGSQWFGRSIHGVKDLGGDLADVAV 598  
 QY 588 VQAQGHVLLLSQPVLRKALMEFNPREVARNVFECNDQVVKVKGKAG-EVRVCLHVQKST 646  
 Db 599 VGAERGVVVLSSRPVDMVTILMSFPABIPVHEVECSYSTSNKMEGVNIITCFQI-KSLY 658  
 QY 647 RDLREGIQSVTVTDALDGRPHSRVAVENETKNSITRQTVLGLTQTCETLKLQLPNC 706  
 Db 659 PQ--FQGRLLANLYTLQDGHMRMSRGLFFDGGSHLSGNTSITP-DKSCLDFFHFFPIC 715  
 QY 707 IEDVPSIVLRLNFSL---VGTPLS-AFGN-----LRPVLAEDAQRFTALFFPEKNCNDITCQ 762  
 Db 716 IQDLISPINVSLNFSLEEBGTPRDQKRAMQPIRLPSIHTV-TKEIPFEKNCKEDKKCE 774



A:Reference number: A45226; MUID:93155124; PMID:8428973

A:Accession: A45226

A>Status: Preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-1151 <BRI>

A:Cross-references: UNIPROT:P56199

A:Experimental source: hepatoblastoma cell line HepG2

A>Note: sequence extracted from NCBI backbone (NCBI:P124326)

F:142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

```

Query Match      18.8%; Score 1103.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 3.3e-68;
Matches 344; Conservative 212; Mismatches 48; Indels 195; Gaps 44;

QY 1 FNLDTENAMTQENARG-FGQSVVQL---QGSVVVVGAPQEIIVAAANORGSLYQCDYSYGS 56
Db 1 FNVDKVSMTPSGPVEDMFGYTVQYQYEEBKGWLLGSLVPGQPKRNTGDKYKCPVGRGE 60

QY 57 CEP-IRLOVPVEA-----VMSLGLSLAAATTPPQLACGPTVHQTCSENTYVKGL 106
Db 61 SLPCVKLDPVNTSIPNVTEVENMTFGSTL-VTPNPGGFLACGPLYAYRCGHLHYTGI 119

QY 107 CFLFGSNLRQPKQPFPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLK- 165
Db 120 CSDVSPFQVNVNSTAP--VQECSTQ-LDIVIVLDCGNSIYPWDSVTA--FLNDLLKMDI 174

QY 166 -KSTFLSLMQYSEEFRIHFTFEKFNPNPRSLIKPITOLLGR-THTATGIRKVVREL 223
Db 175 GPKQTQVGIQYGENVTHEFNLNKYSSTEEVLVAKKIVORGQQTWTALGTDTPARKEAF 234

QY 224 NITNGARKNAFKILITDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPR-----SE 278
Db 235 TEARGARRGVKVMVITVDGESH-DNHRLLKVIQDCEDENIQRFSAIILGSYNRGNLSTE 293

QY 279 KSRQELMTVASKPRDRHVFOINNFALKTIQNLREKIFAIEGTQTCSSSFHEMSQ 338
Db 294 KFVEETKSIASEPTEKFFNFSDDELAVITKTLGERIFALEATADQSAASFEMESQTG 353

QY 339 FSAAITSNGLPLTVSGSYDMAGGVFLYTSKE-----KSTF-INMTVDSMDNDAYLGAA 392
Db 354 FSAHSQDWMLGAVGAYDNWGVVWQKASQIIPNTTNNVSTKKNEL-ASYLGTV 412

QY 393 AIIILNRVQSL-VLGPAPYQHIGLVAMPFRONTGWESNANVKGTQIGAYGASLCSVDVD 451
Db 413 NSATASSGDVLYTAGQPRVNYHTQVLIYRMEDGNIKIQLTSGEQIGSYGSLITTTDID 472

QY 452 SNGSTDVLVLGAPHY-----YEOTR-GGVSVCPVLPGRQARQWQDAVLY 495
Db 473 KSNSTDILLVGAPMYMGTEKEBQGVVYVVALNQTRFQYQMSLEPIKQTCSSRQHNSCTT 532

QY 496 GEQGPWG-RFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAVYVLPHTGSGSISPSHQ 554
Db 533 ENKNEPCGARFGTAIAAVKDLNLDGNDIVIGAPLEDHGGAVIYTHG-SGKTIRKEYAQ 591

QY 555 RTAGSKLSPRLQYFGQSLSGGDLTMDGLVDLTGAGQHVLLRSQVLRVKAIMEFNPR 614
Db 592 RITPSGDGKTLKFFGOSIHGEMDLNGDLTDTVIGLGAALFWSRDAVAVKVTMNFEN 651

QY 615 EVARNVFECNDQVVKGEAG--EVRVCLHVO-KSTFDRLRBGOIQSVVTVYDIALDSGRPH 671
Db 652 KNIQKKNCH---MEGTEVCINATVCFEVKLSKEDTIYADLQ----YRVTLDSLQI 704

QY 672 SRAVFNET-----KNSTRQQTQVLGLTQTCFLKLQLENCTEDPVPVILRLNLSVLT 725
Db 705 SRFSFGTQERKQVORNIYRKSEC-----TKISFYMLDKHFQDSVR-----ITLDFNLT-D 755

QY 726 PLSAFGLRPLVLAEDAQRILTALFPPEKCNKDNIQDDLSITFSPMSLDCLVVGGRPE- 784
Db 756 PENG-----PVLDDSLPNSVHEIYIPAKCKGNEKICISDLSLHVAITEKDLIVRSQNDK 810

QY 785 ENVTVTVRNDGEDSYRTQVTFPPLDLSYRKVSTIQNRQSRQSWRIACBSASTVSGAL 844
Db 811 FNVSLVRKNTKDSAYNTRTIVHYSPLNLFVSGIEAIQK-----SCBSN----- 853

```

```

QY 845 KSTSCSINHPIPPENSEVTENITFDVDSKASLGN-KLLKANVTSENMMRPTNKTEFOLE 903
Db 854 HNTCKVGPFLRRGEMVTFKILFQNTSYLMENVTIYLSATSDSEBPPETILSDNVNALS 913

QY 904 LPVKYAVVMVVTSHGVSTKYLNTFASENTSRVMQHOYQVSN-----LGQR-----L 950
Db 914 IPVKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSPMPDEL 972

QY 951 PISLVF-----LVPRLNQTVIWDRQVTFSENLSSCTCTKE-----RLPS 991
Db 973 KLSISFPNMTSNGVPLYPTGLSS-----SENANCRPHIFEDPFPINSKKWMTT 1021

QY 992 HSDFLAELRKAPVNVCSIAVCQRIQCDIPPFQIOE-----EFNATLK 1033
Db 1022 STD---HLKRGITLDCNTCKPATITCNLTSSDISQVNVSLILMKPTFIKSYFSSINLTIR 1078

QY 1034 GNLSFDVWIKTSHNHLIIVSTABILFNDSVFTLLPGQAFVRSQTETKVEPFEVNPDLPL 1093
Db 1079 GEL-----RSENASLVLSNN-----QKRELATQISKDGLGRVPL 1114

QY 1094 --IVGSSVGGILLALITALYKLGFFPKROYKDMME 1128
Db 1115 WVLLSAFAGLLJLLMLLILALMKLGFKRPLKKMKEX 1151

```

# RESULT 8

I45914

integrin alpha 2 subunit - bovine (fragment)

C:Species: Bos primigenius taurus (cattle)

C>Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 09-Jul-2004

C:Accession: I45914

R:Kamata, T.; Puzon, W.; Takada, Y.

J. Biol. Chem. 269, 9659-9663, 1994

A:Title: Identification of putative ligand binding sites within the I-domain of integrin

A:Reference number: A54402; MUID:94193647; PMID:7511592

A:Accession: I45914

A>Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: mRNA

A:Residues: 1-1170 <KAM>

A:Cross-references: UNIPROT:P53710; GB:I25886; NID:G439695; PIDN:AB59255.1; PID:G439696

F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.5%; Score 1085; DB 2; Length 1170;  
 Best Local Similarity 27.6%; Pred. No. 6.6e-67;  
 Matches 335; Conservative 217; Mismatches 495; Indels 168; Gaps 47;

```

QY 1 ENLDTENAMTFO-ENARGGQSVVQL---QGSVVVVGAPQEIIVAAANORGSLYQCD- 54
Db 19 YNVGLPKAKIFGSPSEQGFYAVQQFINKGNWLLVGSFWSGPPKRMGDDVYKCPVDLST 78

QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPPQLLAGPTVHQTCSENTYVKGLC 107
Db 79 TTCEKLNLTSTSMNSVNTKNTNMSLGLTLTRNVGTGGFLTCGLPWAQCGSQYTTGVC 138

QY 108 FLFGNLRQPKPPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFISTVMEQLK-- 165
Db 139 SDVSPDF-QLRTSPAPAVQTCF-SFIDVVVVCDESNIYPWD--AVKNELEKFXVQGLDIG 194

QY 166 KSKTLFSLMOYSEERIHFTFEKFNPNPRSLIKPITOLL-----GRTHATGIRKVVRE 221
Db 195 PKTQMGTLQIYANPRVFNLTNFKSKD---EMIKATSTQTFQYGGDLTNTFKAIOVARDT 251

QY 222 LFNITNGARKNAFKILITDGEKFGDPLGYEDVPEADREGVIRVIGV-----GDAFR 276
Db 252 AYSTAAGRPGATKVMVVVTDGESH-DGSKLKAVIDQCNKNILRFGIYAVLGNRNALD 310

QY 277 SEKSRQELNVTASKPRDRHVFOINNFALKTIQNLREKIFAIEGTQTCSSSFHEMSQ 336
Db 311 TKNLIKETKAIASIPTEHRHFNVSDEADLLEKAGTIGQIFSIETVQG-GDNFQWMSQ 369

QY 337 EGFSAAIT--SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFPINMT--RVDSDMN-DAYLGYA 391
Db 337 EGFSAAIT--SNGPLLLSTVGSYDMAGGVFLYTSKEKSTFPINMT--RVDSDMN-DAYLGYA 391

```

Db 370 VFSAEYSPQNNILMLGAVGAYDWSVTQVTKTHGHLIFSKQAFBQILQDRNHSSYLIGS 429  
Qy 392 AAILLRNVQSLVLGAPRYQHIGLVAMFRONTGMWENANV-----KGTQIGAYFQASL 445  
Db 430 VASISTGNSVHFVAGAPRANVTGQIVLYSVN-----ENGNVTVQSQRGQIGSYFGSVL 484  
Qy 446 CSVDVDSNGSTDVLIGAPHYYEQTR--GGQSVCPPLRGORARWOCDAVLYGBQGPWG 503  
Db 485 CAVDVNDKTTIDVLLVGAFTMYNDLKKBEGRVYLFITKIG-ILNWH--QFLEGNGLENA 541  
Qy 504 RFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSORLAGS--KL 561  
Db 542 RFGSAIALSDINWDGNDVIVGSPLENQSGAVYIYNGHEGM-IRLRSQKILGSDRAF 600  
Qy 562 SPRLOYFGSLSGQDITMDGLVDLTVGAQGHVLLRSQPVLRVYKALMEFNPREVARNVF 621  
Db 601 SSHLOYFGSLDYGDLNGDSITDVSAGFCQVQLMSQSADVSVDASFPTPKI--TL 658  
Qy 622 ECNDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVYVDLALD---SGRPHSRAVEN 677  
Db 659 NKNAEI-----KLKLCF-----SAKFRPTQNNQVAILVYNTIDEDOFSSRSVIRGLFK 707  
Qy 678 ETKNSTRQTVLGTQFCE--TLKLQPLNCIEDPVSPVLRILNPSL--VGTPLSAFNL 733  
Db 708 ENNERCLOXTMIVSQARCSYVITHIQEPS---DIISPLNLCMISLENPGT-----756  
Qy 734 RPVLAEADAORLFTALFPFEKNCMDNICODDLSTIF-----SPMSLDCLVGGPREFNVTV 789  
Db 757 NPALAEAYSETVKVFSIPPHKCGDDGVCSIDLVLNVQQLPATQQQPFIVSNQNRKLTFSV 816  
Qy 790 TVRNDGESYQTVTFPPFLDLSYRKVSTLQNRQSRWRLACESASST-EVSGALKSTS 848  
Db 817 QLNKKESAYNTEIIVVDFSENLF-----ASWMPVDGTEVTCQIASQKSVT 864  
Qy 849 CSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFFOLELPVKY 908  
Db 865 CNVGPALKSKQVTFITFDNLQ-NLQNASISFRALSQSEENMADNSVNLKLSLY 923  
Qy 909 AVYVMVTVSHGVSTKYLNFTASNTSRVMOHQVQVSNLQOR-----SLPISLVFLV 958  
Db 924 DAEIHT-RSTNINFEVSLDGNVSSV-HSPE--DIGPKFTFSIKVTTGVSFVMSA---976  
Qy 959 PVRLNQTVWRPQVTFSEN--LSSTCHTK-----RLPSHSDFLAE- 998  
Db 977 -----SVLIHTPQYTKOKNPLMYLTGVHTDQAGDISCEAEINPLKIGOTS SVSFKSEN 1030  
Qy 999 LRKAPVWNGSIAVCORIQDIPFGTQBEFNATLKNLSFDWYIKTSHNLLIVSTABI- 1057  
Db 1031 FRHIKELNCRATASCNIMCWLRLDQVKGBYFLNVSTRIWNGTFAASTFTVQLTAAAEID 1090  
Qy 1058 LPNDSVFTL-----LPGGAFVRSQTETKVEPPE-VENPLPLIVGSSVGGILLALITA 1110  
Db 1091 TYNPQIVYIENVTIP-----LTIMPKHEKVEPTGVIIVGSIAGILLALLALVA 1140  
Qy 1111 ALYKLGFFKRYKDM 1125  
Db 1141 ILWKLGFRRKRYEKM 1155

## RESULT 9

S44142  
VIA-2 protein homolog - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S44142  
R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze  
submitted to the EMBL Data Library, January 1994  
A:Description: The mouse VIA-2 homologue supports collagen and laminin adhesion but not  
A:Reference number: S44142  
A:Accession: S44142  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1178 <EDE>

A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:g473098; PTDN:CAA82877.1; PID:g473098  
F;169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.2%; Score 1071; DB 2; Length 1178;  
Best Local Similarity 27.9%; Pred. No. 6.3e-66;  
Matches 342; Conservative 208; Mismatches 487; Indels 188; Gaps 44;

Qy 1 FNLDTENAMTFQ-ENARGFGQSVVQL-----QGSRRVVVGAQOEIIVAANQSGSLYQC--DYST 54  
Db 27 YNVGLPKAKIFSFPSSHQFGYSVQQLTNPQGNLLVGSWSPGFENRMGVDYKCPVDLPT 86  
Qy 55 GSCPIRLQ-----VPVEAVNMISGLSLAATYSPQLLACGTVHQTCSENTYVGLC 107  
Db 87 ATCEKLNQNSASISNVTETKTNMSLGLTLTRNPGTGGFLTCGFLMAHQCGNQYAYGIC 146  
Qy 108 FLGSLNRLOQPO---KPEEARLQCPQEDSLAFIIVDGSGLIIPHDPRRAKEFISTVMQEQL 164  
Db 147 ----SDVSPDFQLTSPSPAVQACPSL-VDVVVVCDSENSIYP--WEAVKNFLVKFVTGL 199  
Qy 165 K--KSKTLFSLMOYSEFRJHFTKFEQNNPNRSLIKPITQLLG-RTHATGTRKVVRE 221  
Db 200 DIGPKTQVALIQYANEPRIIFNLNDFETKEDMVQATSETRQHGGLTNTFRAIEFARDY 259  
Qy 222 LFNITNGARKNAFKILITIDGKFGDPLGVEDVPEADRECVIRYVIGV-----GDAFR 276  
Db 260 AYSQTSGRPCGATKVMVWVTDGESH-DGSKLKTVIQCCNDDEILRFGLAVLGYLNRNALD 318  
Qy 277 SEKSRQELNVTASPRDRHVFIQINNFEALKTIQNLREKIFAIBGTQTGSSSPHEMSQ 336  
Db 319 TKNLIKIKATASFTERYFNFVADAEALKEKAGTLGQIIFSIETVUG-GDNFQWEMAQ 377  
Qy 337 EGSA--AITSNGPLLLSTVSGSYDWAGGVFLYTSKESTFINMT--RVSDMN-DAYLGYA 391  
Db 378 VGFADYAPQNDILMLGAVGAFDWSGTIVQETSHKPVIFPKQAFQDVLQDRNHSFLGYS 437  
Qy 392 AAILLRNVQSLVLGAPRYQHIGLVAMFRONTGMWENANV-----KGTQIGAYFQASLCS 447  
Db 438 VAAISTEDGVHFVAGAPRANVTGQIVLYSVNK--QGNVTVIQSHRGQIGISYFSVLCSS 494  
Qy 448 VDVDSNGSTDVLIGAPHYYEQTR--GGQSVCPPLRGORARWOCDAVLYGBQGPWGFR 505  
Db 495 VDVDDKTTIDVLLGAPTYMNDLKKBEKGVLYFITKILLNQHQ---FLEGEPTGNARF 551  
Qy 506 GAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLFHGTSGSGISPSHSORLAGSLSR- 564  
Db 552 GSAIALSDINWDGNDVIVGSPLENQSGAVYIYNGHGT-IRTKYSQKILGSGAFRR 610  
Qy 565 -LQVFGQSLSGQDITMDGLVDLTVGAQGHVLLRSQPVLRVYKALMEFNPREVARNVFC 623  
Db 611 HLQFFGRSLDYGDLNGDSITDVSIGALGVQVQLMSQSIADVAIEALFTP-----660  
Qy 624 NDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVYVDLALD---SGRPHSRAVENET 679  
Db 661 -DKITLLNKAKITLKLCFRAEFRRAGQNNQV--AILFNMTLDADGHSRVSRTSRGVREN 717  
Qy 680 KNSTRQTVLGLTQTCET--LKLQPLNCIEDPVSPVLRILNFSLVGTPLSAFGNLRPVL 737  
Db 718 SERFLQKNVYVNEVQKCEHHHSIQKPS--DVVNPLDLRVDISLENPGTS-----PAL 768  
Qy 738 AEDAQRLLTALPFPFKNCNDNII CODDLISI-----TFSFMSLDCLVVGGPREFNVT 788  
Db 769 EAYSETVKVFSIPFYKCGSGDGI:SDLILDVQQLPAIQTSF-----IVSNQNRKLTFS 823  
Qy 789 VTVRNDGESYQTVTFPPFLDLSYRKVSTLQNRQSRWRLACESASST-EVSGALKST 847  
Db 824 VILKRGESAYNVTVLAESENLF-----ASFSMPVDGTEVTCVSGSQSV 871  
Qy 848 SC SINHIPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKT-----FQLELP 905  
Db 872 TCDVGVYFALKSEQVTFITFDNLQ-NLQNAAINFQAFSESQ--ETINKADNSVSLTIP 928  
Qy 906 VXYAVYVMVTVSHGVSTKYLNFTASNTSRVMOHQVQVSNLQOR-----SLPISLV 955

```

Db 929 LLYDAELHLT-RSTNINFEISSDENAPSVIK---SVEDIGPKPIFSLKVTAGAPVSM 984
Qy 956 FLV-----PVRNQTVWDRPQVTP-SENLS 980
Db 985 LVTHIPIQYTKKPNLLYLTGIQDQAGDISCTAEINPLKLPHTA---PSVSFKNENFR 1040
Qy 981 STCHTKERLPSHSDFLAELRKAPVWNCISAVCORIQCDIPFGIQEENATLKGNSLSDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMKAEYFINVTVWVRT 1080
Qy 1041 YIKTSHNHLIVSTAEILFNDSVTLIPGQAGFVRSQTETKVEPPEVNPPLPLVGSVVG 1100
Db 1081 FASFTQTVQLTAAAEIDTHNPQLFVEENAVTIPLMIMKPTKAEVPT--GVIIGSIIA 1138
Qy 1101 GLLLALITAAALYKLGFFKQYKDM 1125
Db 1139 GILLLLAMTAGLWKLGFFKQYKDM 1163

RESULT 10
A33998
integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence, revision 18-Sep-1992 #text_change 09-Jul-2004
A;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:P7301; GB:X17033; NID:933906; PIDN:CAA34894.1; PID:933907
A;Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R;Cattell, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 10-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:L24121; NID:9400342; PIDN:AAA16619.2; PID:94583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match
Best Local Similarity 26.9%; Pred. No. 1e-65;
Matches 329; Conservative 214; Mismatches 494; Indels 188; Gaps 43;

Qy 1 FNLDTENAMTFQ-ENARFGQSVQVL---QGSRVVVGAPQEIETVAANQSGSLYQC--DYST 54
Db 30 YNVGLPEAKIFGSPSEQFGYAVQVFINPKGNWLLVGVSPFENRMDGVYTKCPVDLST 89
Qy 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVKGILC 107

```

```

Db 90 ATCEKLNLTSTSPNTVMETKTNMSLGLILTRNMTGGFLTCGLPWAQCCNQYTTGVC 149
Qy 108 FLFGSNLRQOPQKPEALRGCPQEDSDIAFLVDGSGSIIPHDERRAKEFIETVMEQLK-- 165
Db 150 SDISPDF-QLSASFSPATPCPSL-IDVVVCDSESNLYPWD--AVRNFLKFKVQGLDIG 205
Qy 166 KSKTLFLSMQYSEBRFHFTFKFQNNPNRSLIKPITQLLG-RTHATGATGIRKVVRELFN 224
Db 206 PTKTOVGLIOYANNPRVFNLTVYKTEEMIVATSQTSQYGGDLTNTFGALQYARKVAYS 265
Qy 225 ITNGARKNAKILILLIDGKFGDPLGYEDVIEADREGVIRYIGV-----GDAFSEBK 279
Db 266 AASGRRSATYKVVVVVTDGESH-DGSMLKAVIDQCNDHNLTRFGIAGLVGLNRLALDTKN 324
Qy 280 SRQELNTVASKPRDRHVFOINNFEALKTIQNLRKIFAIEGTOTGSSSSFEHMSQEGF 339
Db 325 LIKEIKAIASIPTEYFFNVSDAEALLEKAGTIGEIFSIEGTVOG-GDNFQEMSQVGF 383
Qy 340 SAAITSNGP--LLSTVGSYDWAGGVFLYTSKSKSTFINMT--RVDSDMN-DAYLGYAAAI 394
Db 384 SADYSSQNILLMLGAVGAFGWSGTIVQKTSHGHLIFPKQAFDQILQDRNHSSYLGVSA 443
Qy 395 ILNRNVQSLVLGAPRYOHIGLVAMFRONTGWESNANV-----KGTQIGAYGASLCSV 448
Db 444 ISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNITVIOAHRGDQIGSYFGSLCSV 498
Qy 449 DVDSNGSTDLAVLIGAPHYEQTR--GGQVSVCPPLRGQARWOCDAVLVYEQQCPWGRFG 506
Db 499 DVDKDTITDVLVAGAPWYMSDLKKEGRVYVFIKKGILGQHO---FLEGEIGIENTRFG 555
Qy 507 AALTUVLDVNGDKLTVAIGAPDEENRGAVLPHGTSGSGISPHSQRISAGS--KLSPR 564
Db 556 SAIAALSDINMGDFNDVIVGSPLENQSGAVIYNGHGT-IRTKYSQKILGSDGAFRSH 614
Qy 565 LQYFGOSLGGQDLTWDGLVLTGAGQHVLLRSQVRLVKALMEFNPREVARNPECN 624
Db 615 LQYFGSLDGYGLDNGSDITDVSIGAFQVQVQLWSQSIADVAIEASTPEKI--TLVNKN 672
Qy 625 DQVVKGEAGEVRVCLHVQKSTRDLRREGIQSVVTVYDLALD---SGRPHSAVENETK 680
Db 673 AQII-----LKLCF---SAKFRPKQNNQVAIVYNTILDADGFSRVTSGLFKENN 721
Qy 681 NSTRQTVLGTQTC--ETLKLQLPNCIEDPVSPVLRNLNFSLVGTPLSAFGLRFLVA 738
Db 722 ERLCQKNMVYVNAQSCPEHIIYIQEPS---DVVNSLDLRVDISLENPGTS-----PALE 772
Qy 739 EDAQRFLTALFPPEKMGNDNICQDDLSITF---SFMSLDCLVVGGPREFNVTVVRND 794
Db 773 AYSETAKVPSIFPHKDCGEDGJCISDLVLDVVRQIPAAQEQPFIVSNQKRLTFSVTLKN 832
Qy 795 GEDSYRTQVTFPFLDLSYRKVSTLQNRQSRQSWRIACESASST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVVDFSENLF-----ASPSLPVDGTGTEVTCQVAASQKSVACDVG 880
Qy 854 PIFPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRTNKTQFQLELPVKYAVMV 913
Db 881 PALKREQQVTFITNDFNLQ-NLQQAQSLSFQALSESQEBENKADNLVNLKIPLLYDAEI- 938
Qy 914 VTSHGVSVKYLNFTASENTSRVMOHQYQVSNLQOR-----SLPSLVLV----- 958
Db 939 ---HLTRTNINIFYEISSDGNVPSIVHSFEDVGPKFIFSLKVTTSQVPSMATVHIHPQ 995
Qy 959 -----PVRNQTVWDRPQVTP-SENLSSTCHTKER 988
Db 996 YTKENPLMYLTGVQTDKAGDISCNADINPLKIGQT---SSVSFKSENER---HTKE- 1047
Qy 989 LPSHSDFLAELRKAPVWNCISAVCORIQCDIPFGIQEENATLKGNSLFDWYIKTSHH 1048
Db 1048 -----LNCRTASCSNVTCWLKDVHMKGEYFVNVTVTRIWNGTFASSFTQT 1091
Qy 1049 LLIVSTABI-LFNDVSFTLLPGQAFVRSQTETKVEPEVNP-----LP--LIVGSSVG 1100

```

Db 1092 VQTAABINTYNPRIYVI-----EDNTVTIPLIMKPEDEAEVPTGVIGSIIA 1141

QY 1101 GLLLALITAAALKYKGLFFKQYKDM 1125

Db 1142 GILLALLAVAILWKLGLFFKRYEKM 1166

RESULT 11

A35854

integrin alpha-1 chain precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 23-Oct-1990 #sequence\_revision 13-Sep-1991 #text\_change 09-Jul-2004

C:Accession: A35854; S1243

R:gnatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J. Cell Biol. 111, 709-720, 1990

A:Title: Molecular cloning of the rat integrin alpha-1-subunit: a receptor for laminin A

A:Reference number: A35854; MUID:90338125; PMID:2380249

A:Accession: A35854

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1180 <IGN>

A:Cross-references: UNIPROT:P18614; GB:X52140; NID:956493; PIDN:CAA36384.1; PID:956494

C:Keywords: cell adhesion; cytoskeleton; transmembrane protein

F:170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 18.1%; Score 1062; DB 2; Length 1180;

Best Local Similarity 27.4%; Pred. No. 2.7e-65;

Matches 344; Conservative 201; Mismatches 478; Indels 234; Gaps 47;

QY 1 FNLDENAMTFOENARG-FGQSVVQL---QGSRVVVGAPQEIIVAANORGSLYQCDYSTGS 56

Db 29 FNVVVKMSGSPGVEDMFGYTVQYENEEKWVLGSLVPLVGQPKARTGDYKCVGRER 88

QY 57 CEP-IRLQVPVEA-----VNMSGLSLAAATSPPLLACGPTVHQTQCESTYKVL 106

Db 89 AMPCVKLDLPVNTSIPNVTIKENMTFGSTL-VTPNPGGLACGLPVAIRGCHLHYTTGI 147

QY 107 CFLFGSNLRQOPQKPEALRCPQEDSDIAFLVQSGSIIPHDPRAKEFTSTWEOQLK- 165

Db 148 CSDVSPTFQVNVSEAP--VQECSTQ-LDIVLDGSSNIYP--WESVIAFLNDLLKRWDI 202

QY 166 -KSKTLFLMOYSEFRIFHTFKPEQNNPNRSLIKPITOLLG-RHTTATGIRKVVRELF 223

Db 203 GPKQTQVGIQYGENVTHEFNLNKYSSTEELVVAANKIGRQGLQTMALGIDTARKEAF 262

QY 224 NITNGARKNAFKILILITDGEKFGDPLGVEDVPIPADREGVIRYVIGVGDAPR-----SE 278

Db 263 TEARGARRGVKVMVIVTDGESH-DNYRLKQVIQDCEDENIQRFSAIILGHYNRGNLSTE 321

QY 279 KSRQELNTVASKPPDRDHVQINNFEALKTIONQLEKIFALEGTQTGSSSFHEHMSQEG 338

Db 322 KFVEEIKTASPTSEKHFNFVSDLEALVTIVKALGERIFALEADADQSAASFEMMSQTG 381

QY 339 FSAATITSGPLISTVGSVDWAGVFLYTSKEKSTFINMT--RVDSMDND---AYLGAYAA 393

Db 382 FSAHVSQDWMLGAVGAYDNQNTVMQKQNMVIPHNTTFQTEPAKMWLEPLASLYGTWN 441

QY 394 IILNRVQSLVIGAPRYQHIGLVAMFRONTGMWESNANVKGTOIGAYGCASLCSVDVDSN 453

Db 442 SATIPGDVLYIAGQPRYNHTGQVYIKWMDGNINTLQTLGGHQIGSYFGSVLTTIDIDKD 501

QY 454 GSTDLVLIGAPHY-----YEQTR-GQVSVCPPLPRGQARWOCDAVLGE 497

Db 502 SYTDLLLVGAPMYMTGEEKQKVYVYAVNQTRFQYMSLEPIRQTCSSSLKDNSTKEN 561

QY 498 CQQPWG-RFGAALTVLVDYNGDKLTDVAICAPCEEDNRGAVYLPFGTGGSGTSPSHSQRI 556

Db 562 KNEPCGARGTAIAAKVLDNVDFNDVWIGAPLEDHAGAVIYHG-SGKTIREAYAORI 620

QY 557 AGSKLSPRIQYFGQSLSGGDLTMDGLVDLTVGAQGHVLLLSQPVLRKATMEFNPRV 616

Db 621 PSGGDGKTLKFFGQSIGHGMDLNGDGLTDVTITGGLGGAALFWARDVAVVVKVITNFPENKV 680

QY 617 ARNVFECNDQVVKGEAG--EVRVCLHVQ-KSTRDLRREGQIQSVVYDIALDSGRPHSR 673

Db 681 NIQKNCR---VEGKETVCINATMCFHVKLKSXSDIYBADLQ-----YRVTLDLSRQISR 733

QY 674 AVNET-----KNSTRQTOVLGLTQTCETLKLQPNCI-----EDPVSPIVLR 718

Db 734 SFESGTQERKIQRNITVRESE-----CIRHSFYMLDKHDFQDSVRVTL 776

QY 719 NPSLVGTPLSAFNGNRPVLAEDAQRLLFTALFPPEKCNKNDNIQQDDLSITFSFMSLDCLV 778

Db 777 DFNLT-DPENG-----PVLDDALPNSVHEHIPAKOCGNKERCISDLTLNVSTTEKSLLI 830

QY 779 VGGPRE-FNVTVTRNDGEDSYRTQVTFPPPLDSYRKVSTLQNRQSRQSRWRACESASS 837

Db 831 VKSQHDKFNVLTVKNGDSAYNTRTVVQHSNPLIFSGIEIQKD-----SCSEN-- 880

QY 838 TEVSGALKSTSCSINHPPIFENSEVFNITFDVDSKASLGNKLL-LKANVTSENMPRTN 896

Db 881 -----QNTICRVGVFPFLRAGTIVTFKIIPQNTSHLSNAIHLSATSDSEEPLESIN 933

QY 897 KTEFQLELPVKYAV---YMWVTSHGVSST-----KYLNFTASENTSRVMQHQYQVSNL 945

Db 934 DNEWNTSIPKYEVLGLOFYSSAHEHLSVAANETIPEFINST--EDIGNEINVPYTIKR 991

QY 946 GQRSLP---ISLVP-----LVPRLNQTVIND-----RP-----Q 972

Db 992 GHFMPPELQLSISFPNLTADGYPLVPIG-----WSSSDNVNCRPSRLEDPFGINSKK 1045

QY 973 VTFS-----ENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDI--- 1019

Db 1046 MTISKSVELKRGTIQDCSSTC-----GVATITCSLLPSDLSQVNVSL 1088

QY 1020 ---PFFGIQBEF---NATLKNLSFDWYIKTSHNHLIVSTAETILFNDSVETLLPGQAF 1073

Db 1089 LWKTFP-IRAHFSSINLTGELK-----SENSLTSSN----- 1123

QY 1074 VRSQTEKTVKPEFFVNPPLPL--IVGSSVGGILLIALLTAALYKGLFKPKQYKDMSE 1128

Db 1124 RKRELAIQISKDGLPGRVPLWVILLSAFAGLLMLLILALWKIGFFKRLPKKKMEK 1180

RESULT 12

A41131

Lymphocyte-Peyer's patch adhesion molecule alpha 4 precursor - mouse

N:Alternate names: integrin alpha-4

C:Species: Mus musculus (house mouse)

C:Date: 05-Jun-1992 #sequence\_revision 05-Jun-1992 #text\_change 09-Jul-2004

C:Accession: A41131; S16742

R:Neuhaus, H.; Hu, M.C.T.; Hemler, M.E.; Takada, Y.; Holzman, B.; Weisman, I.L. J. Cell Biol. 115, 1149-1158, 1991

A:Title: Cloning and expression of cDNAs for the alpha subunit of the murine lymphocyte A:Reference number: A41131; MUID:92064645; PMID:1840602

A:Accession: A41131

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1039 <NNU>

A:Cross-references: UNIPROT:Q00651; EMBL:X53176; NID:951484; PIDN:CAA37316.1; PID:951484

C:Superfamily: integrin alpha-4 chain

C:Keywords: cytoskeleton; transmembrane protein

Query Match 11.3%; Score 663; DB 2; Length 1039;

Best Local Similarity 22.7%; Pred. No. 1.4e-37;

Matches 281; Conservative 187; Mismatches 403; Indels 368; Gaps 48;

QY 1 FNLDENAMTFO-ENARGFGQSVV-QLOGSR--VVVGAPQEIIVAAN---QRGSLYQCDY 52

Db 41 YNLDPENALLQPGSGTLFGYSVVLKSHGSKRWLVGAPTASWLSNAAVNVNFGAIVRCGI 100

QY 53 STG---SCEPIQLVP-----VEAVNMSLGLSLAAATSP-PQLLAG---PTVHQ 95

Db 101 RKNPNQTCQLQSGSGEPGCKTCLERDQNLQWGLTSLRQPGENGSIIVTCGRWKNIFY 160

QY 96 TCSNTYVKGCLFLFGSNLRQOPQKPEALRCPQEDSDIAFLVDGSGIIPHDPRAKE 155



Db 161 MKSDNKLPTGICYWPSDLRTELSK----- 185  
QY 156 FISTVMQLKSKTLFSLMQYSBFRHFTFKFQNNPNRSLIKPTQLLGRTHATGI 215  
Db 186 -----RMAPCYKDYT----- 195  
QY 216 RKVVRELFNITNGARKNAFKILLITDGEKFGPLGVEDVIPEADREGVIRYVIGVDAF 275  
Db 196 -----RKFE----- 200  
QY 276 RSEKSRQELNTVASKPRDHVFQINNFEALKTIONQLREKIFAIEGTQTGSSSFEHMS 335  
Db 201 -----NFAS-----C 205  
QY 336 QEGFSAALTSNGPLLSTVSGVDMAGGVFLY--TSKEKSTFTINMTRVSDMNDAYLGYA-- 391  
Db 206 QAGISSFYTDLIIVMGAPGSGYWTGTVFVNIITNYKAFVD--RQNVKFGSYLGYSVG 263  
QY 392 AAILNRVQSLVILGAPRYOHIGLVAMFRQNTGMWESNANV---KGTQIGAYFGASLCS 447  
Db 264 AGHFRSPHTTEVVGAPQHIGKAYIF---SIDENELNIVYEMKKKLGSPFGASVCA 319  
QY 448 VDVSNGSTDLVLIGAPHYETQRTGGQVSCPLPRQRA--RWQCDVLYGEGQGPWGRFG 506  
Db 320 VDLNADGSDL--LVGAPMQSTIREGRVFVY--INSGMGAVVMEMERVLVSGDKYA--ARFG 376  
QY 507 AAILNRVQSLVILGAPRYOHIGLVAMFRQNTGMWESNANV---KGTQIGAYFGASLCS 566  
Db 377 ESTANLGDINDGFEIDAIPQEDDLRGAVIYNGRV--DGISTYSQRIEQQOIKSLR 435  
QY 567 YFQGSLSGGQDLTMDGLDVLTVGA--QGHVLLRSQPLRVKATMEFNPREVARNVFEON 624  
Db 436 MFGQISGQIDADNGVGVAVGAFQSDSAVILRTRPVIVEASLS--HPESVNRTPDCT 494  
QY 625 DQVYKKEAGEVVRVCHLVQKSTRDRREGIOISVTVYDLDLDSGR---PHSRVAVF--NET 679  
Db 495 -----ENGLPSVCMHLTLCFSYKKEVPGYVLFVYNSLVDVHRKAESPFRVFSNGT 547  
QY 680 KNSTRQTVGLTQTCETLKLQPLNCIEPVPSPVILRLNFSL-----VGTPLSAFNLRL 734  
Db 548 SDVITGSIRVSSGGEKCRTHQAPMRKDVDRDILTPHVEATYHGHVITKRNTEEPPLQ 607  
QY 735 PVLAEDAQR--LEFTALPFFKNCNDNICODLSITFS-----FMSLDCLVVGPREFN 787  
Db 608 PILQKKKXKDVIRKMINFAFCAYEN--CSADLVQSAKVGFLKPYENKTYLAVGSMKTIIL 666  
QY 788 TVTVRNDGDSYRTQVTFPPDLVSRKVSTIQNQRQSRWRLACESASSTEVSGALKST 847  
Db 667 NVSLFNAGDDAYETTLNVQLPTGLYPIKILDLLEK-----QINCE--VTSSGIVK-L 716  
QY 848 SCGINHPFPENSEVFNITFDVDSKASLGNKLLKANVTSEN--NMPTNKTEFOLELP 905  
Db 717 ACSLGIYVDRLSRIDISFLVSSLSRAHEDLSISVHASCENEGELDQVRNVRVLTIP 776  
QY 906 VKYAVVMVTSHG--STKYLNTFASNTSRVMQHQ-----YQVNLGQSLP--ISLVFL 957  
Db 777 LAYEV--MLTVHGLVNPTSFVYGSSENEPETCMAEKLNIITHIVNTGISMAPNVSVKIM 834  
QY 958 VPVRLNQTIVWRP--QVTFSENLSSTCHTKE-----RLPSSHDFLAELR 1000  
Db 835 VP---NSFLPQDDKLFLNVDVQTTTCQCHFKYHGRECTFAQQGKIAGTLTDIVKFLSKTD 891  
QY 1001 KAPVNVCSIA--VCQRIQCDIPFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEL 1058  
Db 892 KR--LLYCMADQHCULDFLCN---FGWESG-----KEASVHIQLEGRPSIL 933  
QY 1059 FNDVSFTLLPGQAFVRSQETKVEFVFPNP----- 1090  
Db 934 ENDETSSL-----KFEIKATAFPHPKVIENLKNDENVAHVLEGLHHQRPKRHF 983  
QY 1091 -LPLVIGSVGGILLIALIALTAALYKLGFFKQYKMMSE 1128  
Db 934 ENDETSSL-----KFEIKATAFPHPKVIENLKNDENVAHVLEGLHHQRPKRHF 983

Db 984 TIIITITSLILGLLIVLLISCVMMKAGFFRQYKSILOE 1022  
RESULT 13  
S06046  
Integrin alpha-4 chain precursor - human  
N:Alternate names: CD49d; very late antigen-4 alpha chain; VLA-4 alpha chain  
C:Species: Homo sapiens (man)  
C>Date: 01-Dec-1989 #sequence revision 01-Dec-1989 #text\_change 09-Jul-2004  
C:Accession: S06046; A39355; D28018  
R:Takada, Y.; Ellices, M.J.; Crouse, C.; Hemler, M.E.  
EMBO J. 8, 1361-1368, 1989  
A:Title: The primary structure of the alpha(4) subunit of VLA-4: homology to other integrins  
A:Reference number: S06046; MUID:89356603; PMID:2788572  
A:Accession: S06046  
A:Molecule type: mRNA  
A:Residues: 1-1038 <TA2>  
A:Cross-references: UNIPROT:P13612; GB:X16983; EMBL:X15356; NID:g33945; PIDN:CAA34852.1.  
R:Rosen, G.D.; Birkenmeier, T.M.; Dean, D.C.  
Proc. Natl. Acad. Sci. U.S.A. 88, 4094-4098, 1991  
A:Title: Characterization of the alpha-4 integrin gene promoter.  
A:Reference number: A39355; MUID:91239513; PMID:2034655  
A:Accession: A39355  
A:Molecule type: DNA  
A:Residues: 1-93 <ROS>  
A:Cross-references: GB:M62841  
R:Takada, Y.; Strominger, J.L.; Hemler, M.E.  
Proc. Natl. Acad. Sci. U.S.A. 84, 3239-3243, 1987  
A:Title: The very late antigen family of heterodimers is part of a superfamily of molecules  
A:Reference number: A94151; MUID:87204112; PMID:3033641  
A:Accession: D28018  
A:Molecule type: protein  
A:Residues: 40-50, 'E', 52-53 <TA2>  
C:Genetics:  
A:Gene: GDB:ITGA4; CD49D  
A:Cross-references: GDB:128032; OMIM:192975  
A:Map position: 2q31-2q32  
C:Superfamily: integrin alpha-4 chain  
C:Keywords: cell adhesion; cytoskeleton; duplication; glycoprotein; heterodimer; transmembrane  
F:1-39/Domain: signal sequence #status predicted <SIG>  
F:40-1038/Product: integrin VLA-4 alpha-4 chain #status predicted <MAT>  
Query Match 10.8%; Score 635; DB 2; Length 1038;  
Best Local Similarity 25.3%; Pred. No. 1.2e-35;  
Matches 249; Conservative 154; Mismatches 370; Indels 212; Gaps 39;  
QY 272 GDAFRSEKSRQELNTVASKPRD-----HVFQINNFEALKT-----IQN 310  
Db 121 GKTCLLEERDNQWLGVTLSPQGENSIVTCGHRWKNIIFYIKNENKLTGCGYGVPPDLRT 180  
QY 311 QIREKI-----FAIGTQTGSSSFEHMSQEGFSAAITSGNGLPLSTVGSY 356  
Db 181 ELSKRIAPCYQDYVKKFGENFA-----SCQAGISSFYTKDLIVMGAPGSS 225  
QY 357 DWAGGVFLY--TSKEKSTFTINMTRVSDMNDAYLGYA--AAILNRVQSLVILGAPRYQ 411  
Db 226 YMTGSLFVNIITNYKAPLDKQNVKF---GSYLGSYVAGHFRSQHTTEVVGAPQHE 282  
QY 412 HIGLVAMFRQNTGMWESNANV---KGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHY 467  
Db 283 QIGKAYIF---SIDEKELNILEHMKKLGISFGASCVAVDLNDAGFSDL--LVGAPMQS 337  
QY 468 EOTRGQVSVCPPLPRGQARMQC--DAVLYGEGQGPWGRFGAALTVLGVGVNGDKLTDVAIG 526  
Db 338 TIREGRVFVY--INSGMGAVVMEMERVLVSGDKYA--ARFGESIVNLGIDINDGFEVDAIG 395  
QY 527 APGEEDNAGAVLFGHTSGSGISPSHSQRIAGSKLSPRLVQFQSLSGGQDLTMDGLVDL 586  
Db 396 APQEDDLQAGIYIYNGRA--DGISTFSQRIEGLQISLSLMSFGQISQSIDADNNGYVDV 454  
QY 587 TVGA--QGHVLLRSQPLRVKATMEFNPREVARNVFEONQVVKKEAGEVVRVCHLVQK 644  
Db 455 AVGAFRSDSAVLLRTRPVVIVDASLS--HPESVNRTKFDC-----VENGWSPSCVLDLTL 506





A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1041 <R>  
A;Cross-references: UNIPROT.O76378; EMBL.AF067658; NID:g3202040; PID:g3202041; PIDN:AAC2  
A;Experimental source: developmental stage embryo  
C;Function:  
A;Description: binds laminin  
C;Superfamily: integrin alpha-2b chain

Query Match	10.5%;	Score 614.5;	DB 2;	Length 1041;
Best Local Similarity	24.8%;	Pred. No. 3.3e-34;		
Matches 236;	Conservative 155;	Mismatches 314;	Indels 245;	Gaps 41;

QY	336	QEGSAAITSNCP--LLSTVGSYDWAGGVFLYTSKEKSTFINMTWDS-----DMNDAYLG 388
Db	181	QAGSGIIFSNALSALVMPGASGYLQGOYVQSILNRSV-VQATQESNTGTYSFDSNYSRG 239
QY	390	YAAAI--ILNRVQSVILGAPRYOHI-GIVAMFRONTGHWESNANVKGTQICAYGASLC 446
Db	240	YSLALGDFNGGVQDYVVGTPRAESLMLGVAIFDOQLNQFN---QVMGTQIVAYFGYSVT 296
QY	447	SVVDNSGSTDLVLTIGAPHYEQTRGGGVQSVCLPRGORARWOCDAVLYGEO-----498
Db	297	VVDI-NNDTYDLLVGAPYMWDDGPAIQ-----RWEAGAVYYIQNDPOVGPCA 343
QY	499	-----GQPMWFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAVLFGHTSGS 546
Db	344	SNRLSSTLLTGGQIRGRFGLSIASIGDSNQDGFNDVAIGAPYEGDAGAVVIHG-SAN 402
QY	547	GISPSHSQRIAGKLS-PRLOYFGQSLSGGDLTMOGLVDLTVGAAQ---GHVLLRSQVPL 603
Db	403	GLKSTPAQVLPTSLTGHSGITTFGSLGGQDMDKNYPDILLVGAESANTAVLIRTPVW 462
QY	604	RVKAIMEFNPREVARNPEFCNDQVVVKGEAGEVRVCLHVOKSTRDRLEGQIQSV-----658
Db	463	SLDATLANTEP-----IGINLEKTYE-LADGTMVTSFIAMT 497
QY	659	-----VTYDLALDSG-RPHSRVAFNETKNSTRROTQVLGL-TQTCETLKIQ 702
Db	498	CFTVTGNYPDHIDISYTVTVDSGLIANRRAMFVDNDMSEITKTRLAVSTQFCPLRAY 557
QY	703	LPNCIEDPSPIVURLNPSLVGTPLSAPGN-----LRPVLAEADAQRILFTALFFPEK 753
Db	558	VGNSIEDKLTPIKVTLOYDL-----NNDESRLQPHEILPIIDMATMSTQTKQVSQN 609
QY	754	NCGNNDTCQDDLSTTFSPMSLDCLVGCGPREFNVTVTRNDGEDSYRQTQVTFPPPLDSY 813
Db	610	NCVN-NICIPDLDTVT-PNLPNVIGQTBELTDVSLNNRGEAFOSLSVYVPLGLQF 667
QY	814	RKYSTLQNRQSRSWRLACBSASTSVSGALKJSTCSINHPIFFEN-----SEVTENIT 867
Db	668	VRL-----ERKANMDFSVCSEDS-----LRITCTGNPMWCKNILEFGLJLSTQVS 717
QY	868	FDVDS-----KASIGNKLLKANVTSENNPRNTKTEFQLELPVKYAVYVWVTHSGVSTK 922
Db	718	GDKSDSIEFFYKAEBSNS--EDPNTLENNELMNTVPVTDCTLKLSASYPIVMYSTQED 775
QY	923	YL-----NFTASENISRVMQHOYQV-----SNLGORSUPIS-----LVFLVPV 960
Db	776	YVPPPPPAKNAEADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPKNDEGEYLFYLLGI 834
QY	961	RLNQTVIWDRPQ-----VTFSENLSSCTHKERLPSHSDFLAELRKAPV 1005
Db	835	MTEEGVTCOLTQGANPEGVKLEPSTKAKLSNSTQVSGRKRREPEVAEALAQTDN--VI 892
QY	1006	NCSIAVCQRIQCDIPFGIOEFENAT-----LKGNLSFDWIKTSHNHLILIVSTAEILF 1059
Db	893	YCASDSVCLINCTI-----DEINASKSVVRILGRF--W-----ERTF 928
QY	1060	NDSVFYLLPGQAFVRSQTEKVE--PFEVDPN-----LP-----1092
Db	929	QKAVSELTPVQVQATIASASAANKTIPIYNTPLPRDFSQDSTKASTLVTEELVPPVPIAW 988

Qy	1093	-LTVSSVGGILLALITAALKYKLGFFKR-----QYKDMS--EGGPP 1132 : : :
Db	989	WIIIVSVLGGIILLLIILGLWKCGFFERKKPGEKEYPVASADKGPP 1038 : : :

Search completed: November 9, 2004, 12:19:34  
Job time : 33.25 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds  
(without alignments)  
4783.919 Million cell updates/sec

Title: US-09-902-481B-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTQENARGFGQ.....FKRQYKDMSEGGPPGAEPPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5846.5	99.5	1152	1 ITAM_HUMAN	P11215 homo sapien
2	4470	76.1	1153	1 ITAM_MOUSE	P05555 mus sapien
3	4381	74.6	1151	2 Q9J130	Q9J130 rattus norv
4	3809.5	64.8	920	2 Q28984	Q28984 sus scrofa
5	3479	59.2	1163	1 ITAH_HUMAN	P20702 homo sapien
6	3411	58.0	1162	1 ITAD_HUMAN	P20702 homo sapien
7	3329.5	56.7	1188	2 Q6KAS4	Q6KAS4 mus musculus
8	3329.5	56.7	1188	2 BAD21383	BAD21383 mus musculus
9	3320.5	56.5	1169	1 ITAX_MOUSE	Q9QXH4 mus musculus
10	3244.5	55.2	1161	1 ITAD_RAT	Q9QYF7 rattus norv
11	1583	26.9	1165	1 ITAL_BOVIN	P61625 bos taurus
12	1557.5	26.5	1170	1 ITAL_HUMAN	P20701 homo sapien
13	1557	26.5	1166	2 Q6TYB8	Q6TYB8 bos taurus
14	1557	26.5	1166	2 AAQ90015	AAQ90015 bos taurus
15	1538.5	26.2	1163	1 ITAL_MOUSE	P24063 mus musculus
16	1534.5	26.1	1161	2 Q9RTV4	Q9RTV4 mus musculus
17	1524	25.9	1160	2 Q9RT200	Q9RT200 mus musculus
18	1409	24.0	1196	2 Q98TF1	Q98TF1 cyprinus ca
19	1359.5	23.1	1086	2 Q96HB1	Q96HB1 homo sapien
20	1358.5	23.1	1187	2 Q98TF0	Q98TF0 cyprinus ca
21	1278	21.7	927	2 Q8HZV0	Q8HZV0 bos taurus
22	1167.5	19.9	1167	2 Q88340	Q88340 rattus norv
23	1161.5	19.8	1167	1 ITAE_MOUSE	P38570 mus musculus
24	1153	19.6	1179	1 ITAE_HUMAN	P38570 mus musculus
25	1123	19.1	1167	2 Q88341	Q88341 rattus norv
26	1103.5	18.8	1151	1 ITAL_HUMAN	P56199 homo sapien
27	1093.5	18.6	1189	1 ITAH_HUMAN	Q9UKX5 homo sapien
28	1085	18.5	1170	1 ITA2_BOVIN	P53710 bos taurus
29	1071	18.2	1178	1 ITA2_MOUSE	Q62469 mus musculus
30	1071	18.2	1178	2 Q6P1C7	Q6P1C7 mus musculus
31	1071	18.2	1178	2 AAH65139	AAH65139 mus muscu

## RESULT 1

ID	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
AC	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor).			
GN	Name=ITCAM; Synonyms=CR3A, CD11B;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaout M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mo1: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaout M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mo1 (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098893; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			

## ALIGNMENTS



FT	STGNAL	1	16	Integrin alpha-M.
FT	CHAIN	17	1152	
Query Match 99.5%; Score 5846.5; DB 1; Length 1152;				
Best Local Similarity 99.2%; Pred. No. 0;				
Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;				
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAVQSGSLYQCDYSTGSCPEI	60	
DB	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAVQSGSLYQCDYSTGSCPEI	76	
QY	61	RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGCLFGLGSLNRQOPQK	120	
DB	77	RLQVPVAVNMSLGLSLAATTSPPQLACGPTVHQTCSENTYVKGCLFGLGSLNRQOPQK	136	
QY	121	FPALRGCPDEDSDIAPLVDSGSIIPDPRRAKEPSTVMEQLKSKTILFSLMQYSEEF	180	
DB	137	FPALRGCPDEDSDIAPLVDSGSIIPDPRRAKEPSTVMEQLKSKTILFSLMQYSEEF	196	
QY	181	RIHFTFKFQNNPNRSLIKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI	240	
DB	197	RIHFTFKFQNNPNRSLIKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILIVI	256	
QY	241	TDGEKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN	300	
DB	257	TDGEKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNTVASKPPRDHVFQIN	316	
QY	301	NFEALATIQNLREKIPALEGTGTGSSSPHEMSQEGFSAATISNGPLLSVGVSDWAG	360	
DB	317	NFEALATIQNLREKIPALEGTGTGSSSPHEMSQEGFSAATISNGPLLSVGVSDWAG	376	
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALILNRNVQSLVGLAPRYQHILGVAMFR	420	
DB	377	GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAALILNRNVQSLVGLAPRYQHILGVAMFR	436	
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEYQTRGGQVSCPL	480	
DB	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEYQTRGGQVSCPL	496	
QY	481	PRGORARWCCDAVLXGEOCPGREGAALTGLVDVNGDKLTDVAIGAPGEEDNRGAVYLF	540	
DB	497	PRG-RARWCCDAVLXGEOCPGREGAALTGLVDVNGDKLTDVAIGAPGEEDNRGAVYLF	555	
QY	541	HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGDLTMDGLVDITVGAQGHVLLRSQ	600	
DB	556	HGTSGSGISPSHSORLAGSKLSPRLQYFGQSLGGDLTMDGLVDITVGAQGHVLLRSQ	615	
QY	601	PVLRVKAIMFEPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	660	
DB	616	PVLRVKAIMFEPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVT	675	
QY	661	YDLALDSGRPHGRAVENETKSTRQTQVLGLTQTCETIKLQLPNCIEDPVPVILRLNF	720	
DB	676	YDLALDSGRPHGRAVENETKSTRQTQVLGLTQTCETIKLQLPNCIEDPVPVILRLNF	735	
QY	721	SLVGTPLSAFGLNLRPLVAEDAQRLLTALPFPFKNCNDNICDDLSITPFSMSLDCVLVG	780	
DB	736	SLVGTPLSAFGLNLRPLVAEDAQRLLTALPFPFKNCNDNICDDLSITPFSMSLDCVLVG	795	
QY	781	GPREFNVTVVNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRQSWLACESASSTEV	840	
DB	796	GPREFNVTVVNDGEDSVRTQVTFPPFLDLSYRKVSTLQNRQSRQSWLACESASSTEV	855	
QY	841	SGALKSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRNKTEF	900	
DB	856	SGALKSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRNKTEF	915	
QY	901	QLELPKVAVYVMTVSHGVS TKYLNFTASENTSRVWQHOYQVSNLQORSILPISLFLPVP	960	
DB	916	QLELPKVAVYVMTVSHGVS TKYLNFTASENTSRVWQHOYQVSNLQORSILPISLFLPVP	975	
QY	961	RLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHDFLAELRKAPVNCISIAVCQRIQCIP	1020	

Db	976	RLNQTVIWRDPQVTFSENLSSTCHTKERLPKSHDFLAELRKAPVNCISIAVCQRIQCIP	1035
QY	1021	FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQET	1080
Db	1036	FFGIQBEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQET	1095
QY	1081	KVEPFEPVNPPLPLIVGSSVGGMLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ	1137
Db	1096	KVEPFEPVNPPLPLIVGSSVGGMLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ	1152

## RESULT 2

ID	ITAM_MOUSE	STANDARD;	PRT;	1153 AA.
AC	P05555; Q8CA73;			
DT	01-NOV-1988 (Rel. 09, Created)			
DT	01-FEB-1991 (Rel. 17, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor MOI).			
GN	Name=Itgam;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RX	MEDLINE=88312584; PubMed=3044779;			
RA	Pytela R.;			
RT	"Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the integrin family and an additional domain related to von Willebrand factor."			
RL	EMBO J. 7:1371-1378(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A. (ISOFORM 2).			
RC	STRAIN=CS7BL/6J; TISSUE=Spinal cord;			
RX	MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;			
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,			
RA	Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,			
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,			
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,			
RA	Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,			
RA	Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,			
RA	Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,			
RA	Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,			
RA	Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,			
RA	Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,			
RA	Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,			
RA	Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,			
RA	Nagashima T., Numata K., Okido T., Pavan W.J., Perlea G., Pesole G.,			
RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,			
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,			
RA	Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,			
RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,			
RA	Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,			
RA	Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,			
RA	Tuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,			
RA	Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,			
RA	Shiaki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,			
RA	Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,			
RA	Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,			
RA	Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,			
RA	Birney E., Hayashizaki Y.;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs."			
RL	Nature 420:563-573 (2002).			
RN	[3]			
RP	SEQUENCE OF 11-45 FROM N.A.			
RC	STRAIN=BALE/c; TISSUE=Spleen;			
RX	MEDLINE=86287312; PubMed=2942940;			
RA	Sastre L., Roman J.M., Teplow D.B., Dreyer W.J., Gee C.E.,			
RA	Larson R.S., Roberts T.M., Springer T.A.;			
RT	"A partial genomic DNA clone for the alpha subunit of the mouse			

RT complement receptor type 3 and cellular adhesion molecule Mac-1-11;

RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648(1986).

RP [4]

RA MEDLINE=85188276; PubMed=3887182;

RX Springer T.A., Teplov D.B., Dreyer W.J.;

RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion

RT glycoproteins and unexpected relation to leukocyte interferon.";

RL Nature 314:540-542(1985).

CC -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various

CC adhesive interactions of monocytes, macrophages and granulocytes

CC as well as in mediating the uptake of complement-coated particles.

CC It is identical with CR-3, the receptor for the iC3b fragment of

CC the third complement component. It probably recognizes the R-G-D

CC peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for

CC fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides

CC of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in

CC mast cell development and in immune complex-mediated

CC glomerulonephritis. Mice expressing a null mutation of the alpha-M

CC subunit gene demonstrate increase in neutrophil accumulation, in

CC response to a impaired degranulation and phagocytosis, events that

CC apparently accelerate apoptosis in neutrophils. These mice develop

CC obesity.

CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M

CC associates with beta-2.

CC -!- SUBCELLULAR LOCATION: Type I membrane protein.

CC -!- ALTERNATIVE PRODUCTS:

CC Name=1; Event=Alternative splicing; Named isoforms=2;

CC Name=2; IsoId=P05555-1; Sequence=Displayed;

CC Name=2; IsoId=P05555-2; Sequence=VSP\_010473;

CC Note=No experimental confirmation available;

CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and

CC granulocytes.

CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins

CC with I-domains do not undergo protease cleavage.

CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

CC -!- SIMILARITY: Contains 1 VWFA domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; X07640; CAA30479.1; -;

DR EMBL; AK039444; BAC30350.1; -;

DR EMBL; M14293; AAA39484.1; -;

DR PIR; S00551; S00551.

DR HSSP; P11215; 1BHQ.

DR MGI; MGI:96607; Itgam.

DR GO; GO:0009897; C:external side of plasma membrane; IDA.

DR GO; GO:0007155; P:cell adhesion; IMP.

DR GO; GO:0045123; P:cellular extravasation; IMP.

DR GO; GO:0030593; P:neutrophil chemotaxis; IMP.

DR InterPro; IPR000413; Integrin\_alpha.

DR InterPro; IPR02035; VWF\_A.

DR Pfam; PF01839; FG-GAP; 3-

DR Pfam; PF00357; Integrin\_alpha; 1.

DR Pfam; PF00092; VWA; 1.

DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.

DR SMART; SM00191; Int\_alpha; 5.

DR SMART; SM00327; VWA; 1.

DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.

DR PROSITE; PS02334; VWA; 1.

KW Alternative splicing; Calcium; Cell adhesion;

KW Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;

KW Signal; Transmembrane.

FT	SIGNAL	1	16	Integrin alpha-M.
FT	CHAIN	17	1153	Extracellular (Potential).
FT	DOMAIN	17	1105	Potential.
FT	TRANSNEM	1106	1129	Cytoplasmic (Potential).
FT	DOMAIN	1130	1153	FG-GAP 1.
FT	REPEAT	31	84	FG-GAP 2.
FT	REPEAT	85	163	VWFA.
FT	DOMAIN	164	350	FG-GAP 3.
FT	REPEAT	337	400	FG-GAP 4.
FT	REPEAT	401	452	FG-GAP 5.
FT	REPEAT	454	515	FG-GAP 6.
FT	REPEAT	517	575	FG-GAP 7.
FT	REPEAT	580	632	Potential.
FT	CA_BIND	465	473	Potential.
FT	CA_BIND	529	537	Potential.
FT	CA_BIND	592	600	GFPR motif.
FT	SITE	1132	1136	By similarity.
FT	DISULFID	66	73	By similarity.
FT	DISULFID	105	123	By similarity.
FT	DISULFID	654	711	By similarity.
FT	DISULFID	770	776	By similarity.
FT	DISULFID	999	1023	By similarity.
FT	DISULFID	1028	1033	By similarity.
FT	CARBOHYD	58	58	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	86	86	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	391	391	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	734	734	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	772	772	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	801	801	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	881	881	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	907	907	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	941	941	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	980	980	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	994	994	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1022	1022	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1045	1045	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1051	1051	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	1076	1076	N-linked (GlcNAc. .) (Potential).
FT	VARSPIC	453	569	Missing (in isoform 2).
FT	CONFLICT	37	37	/FtId=VSP_010473.
FT	CONFLICT	683	683	N -> S (in Ref. 2).
FT	CONFLICT	1153	1153	V -> G (in Ref. 2).
FT	SEQUENCE	1153	127480	MW: 178DB988AECB0343 CRC64;

Query Match 76.18; Score 4470; DB 1; Length 1153;

Best Local Similarity 73.94; Pred. No. 7.9e-287; Indels 2; Gaps 2;

Matches 841; Conservative 146; Mismatches 149;

Qy	1	FNLDTENAMTFQENARGGQSVVQLQGSRRVVVVGAPQIVAAQNRGSLYQCDYTGSCPEI	60
Qy <td>17</td> <td>FNLDTENAMTFQENARGGQSVVQLQGSRRVVVVGAPQIVAAQNRGSLYQCDYTGSCPEI</td> <td>76</td>	17	FNLDTENAMTFQENARGGQSVVQLQGSRRVVVVGAPQIVAAQNRGSLYQCDYTGSCPEI	76
Qy <td>61</td> <td>RLQVPVAVNMVSLGLSLAATTSPPQLACGPTVHTQTSNTYVVKGLCFGLFSGNLRQPOK</td> <td>120</td>	61	RLQVPVAVNMVSLGLSLAATTSPPQLACGPTVHTQTSNTYVVKGLCFGLFSGNLRQPOK	120
Qy <td>77</td> <td>PLQVPVAVNMVSLGLSLAATTSPPQLACGPTVHTQTSNTYVVKGLCFGLFSGNLRQPOK</td> <td>136</td>	77	PLQVPVAVNMVSLGLSLAATTSPPQLACGPTVHTQTSNTYVVKGLCFGLFSGNLRQPOK	136
Qy <td>121</td> <td>FPEALRCQPEDSDIAFLVDGSGSIIIPDFERRAKETSTVMEQLKSKTLFSLMOYSEEF</td> <td>180</td>	121	FPEALRCQPEDSDIAFLVDGSGSIIIPDFERRAKETSTVMEQLKSKTLFSLMOYSEEF	180
Qy <td>137</td> <td>FPEALRCQPEDSDIAFLVDGSGSIIIPDFERRAKETSTVMEQLKSKTLFSLMOYSEEF</td> <td>196</td>	137	FPEALRCQPEDSDIAFLVDGSGSIIIPDFERRAKETSTVMEQLKSKTLFSLMOYSEEF	196
Qy <td>181</td> <td>RIHFTTFKEFQNNPNRSLIKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI</td> <td>240</td>	181	RIHFTTFKEFQNNPNRSLIKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI	240
Qy <td>197</td> <td>RIHFTTFKEFQNNPNRSLIKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI</td> <td>256</td>	197	RIHFTTFKEFQNNPNRSLIKPIITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI	256
Qy <td>241</td> <td>TDGEKFGDPLGYEDVIEADREGVIRVIVGVGDFAFRSEKSOELNTVASKPRDHFVFOIN</td> <td>300</td>	241	TDGEKFGDPLGYEDVIEADREGVIRVIVGVGDFAFRSEKSOELNTVASKPRDHFVFOIN	300
Qy <td>257</td> <td>TDGEKFGDPLGYEDVIEADREGVIRVIVGVGDFAFRSEKSOELNTVASKPRDHFVFOIN</td> <td>316</td>	257	TDGEKFGDPLGYEDVIEADREGVIRVIVGVGDFAFRSEKSOELNTVASKPRDHFVFOIN	316
Qy <td>301</td> <td>NFEALKTIONCLREKIFAIEGTQTGSSSPFHEMSQEGFSAAITSNGLLSTVGSYDWAG</td> <td>360</td>	301	NFEALKTIONCLREKIFAIEGTQTGSSSPFHEMSQEGFSAAITSNGLLSTVGSYDWAG	360
Qy <td>317</td> <td>NFEALKTIONCLREKIFAIEGTQTGSSSPFHEMSQEGFSAAITSNGLLSTVGSYDWAG</td> <td>376</td>	317	NFEALKTIONCLREKIFAIEGTQTGSSSPFHEMSQEGFSAAITSNGLLSTVGSYDWAG	376



QY 361 GVFLYTSKEKSTFTNMTVRDSDMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420  
 Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLGASAVILNRVQSLVLGAPRYQHIGLVAMFR 436  
 QY 421 QNTGMWESNANVKTQTGAIFGASLCSVDVDSNGSTDLVLGAPHYETQRTGGQVSCPL 480  
 Db 437 ENFTWEPHTSISKQSGSYFGASLCSVDMDADGNTNLILGAPHYETKRTGGQVSCPL 496  
 QY 481 PRGORARWQCDVLYGQGOQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
 Db 497 PRG-RARWQCEALHGDQHPGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYIF 555  
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 600  
 Db 556 YGASIASLSASHSHRIIAGHFSGLQYFGQSLGGKDLTMDGLMDLAVGAGQHLLLRQA 615  
 QY 601 PVLRVKAIMENPREVARNVFCNDQVVKGEAGEVRCVLHVQKSTRDRLRREGIOQSVVT 660  
 Db 616 PVLRLIATMBFSPKKVARSVFACQVLRKNDAGEVRCVLRVKNTRDRLRREGDIQSTVT 675  
 QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
 Db 676 YDLALDPGRSVVRAPFDETKNGILRRIRVFGTQKCETLKLIPDCVMSVSPITLRNY 735  
 QY 721 SLVGTPLSAFNLRLPVLAEADAORLTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
 Db 736 TLVGEPLRSFNLRLPVLADQAFRTAMFPFEKNCNDNICQDDLSITMSAMGLDILVVG 795  
 QY 781 GPREENVTVRDGEDSYTQVTFPPDLSTVRKYSTLQNRQSRWSRL-ACESASSTE 839  
 Db 796 GPQDFNMSVTLRNDGEDSYTQVTVYVPSGLSVRKDSASQNPULTKKPWFVKPAESSSSE 855  
 QY 840 VSGALSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSNNPRTNKT 899  
 Db 856 GHGALSTTWINHPFPANSEVTFNITFDVDSHAFGNKLLKAIIVASENNMSRTHKT 915  
 QY 900 FOLELPKVAIVMVTSHGVSTKYLNTFASNTSRVQHQYQVSNLQORSLPISLVFLVP 959  
 Db 916 FOLELPKVAIVMVTDESSIRYLNFTASEMTSKVIHQYQVNNLQORSLPVSVVFWIP 975  
 QY 960 VRLNQTVIWRDPQVTSNLSSTCHTKERLPSHSDFLAEIRKAPVNVNCSIAVQRIQCDI 1019  
 Db 976 VQNNVTVDHPQVIFSONLSSACHTEQSPHSNFRDQLERTPVLNCSVAVCKRIQCDL 1035  
 QY 1020 PFGIOEEFNATLKNLSPDWYIKTSHNLLIYSTAEILFNDSVFTLLPQCGAFVRSQTE 1079  
 Db 1036 PSNTQEIFNVILKGNLSPDWYIKTSHGHLNLSSTAEILFNDSAFALLPQGESVRSKTE 1095  
 QY 1080 TKVERFEVNPPLIIVGSSVGGLLLALITALIYKLGFFKRYQKDMWSEGGPPGAPQ 1137  
 Db 1096 TKVERFEVNPPLIIVGSSIGGLVLLALITAGLYKLGFFKRYQKDMWNEAAPQDAPQ 1153

RESULT 3  
 Q9UI30  
 ID Q9UI30 PRELIMINARY; PRT; 1151 AA.  
 AC Q9UI30;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Integrin beta 2 alpha subunit.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fathallah D.M. Sr., Zeria K. Jr.,  
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 DR EMBL; AF268593; AAF81280.1; -  
 DR HSSP; P11215; 1BHQ.

DR GO: 0008305; C:integrin complex; IEA.  
 DR GO: 0007160; P:cell-matrix adhesion; IEA.  
 DR GO: 0007229; P:integrin-mediated signaling pathway; IEA.  
 DR InterPro; IPR00413; Integrin\_alpha.  
 DR InterPro; IPR02035; VWF A.  
 DR Pfam; PF01839; FG-GAP; 2.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00032; VWA; 1.  
 DR PRINTS; PRO1185; INTEGRIN.  
 DR PRINTS; PRO0453; VWFADOMAIN.  
 DR SMART; SM00191; Int alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS0234; VWF A; 1.  
 DR Cell adhesion; Integrin; Transmembrane.  
 KW

SQ SEQUENCE 1151 AA; 126943 MW; 8F785695D4074CA5 CRC64;

Query Match 74.6%; Score 4381; DB 2; Length 1151;  
 Best Local Similarity 72.7%; Pred. No. 6.2e-281;  
 Matches 827; Conservative 150; Mismatches 158; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGSLYQCDYSTGSCPEI 60  
 Db 17 FNLDTENPMTFQENASGFGQSVIQLGETRVVAAAPQEVKAVNQTGALYQCDYSTNCRDPI 76  
 QY 61 RLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHQTCSENYYVKGICFLFGSNLRQOPQK 120  
 Db 77 PLQVPEAVNMSLGLSLAATTVPQALLACGPTVHQNCENTYVNGLCYLFSGNLLRQPOQ 136  
 QY 121 FPEALRGCEQEDSDIAPLVDSGSIIPHFRRAKERISTVMEOLKSKSTLFSIMQYSEEF 180  
 Db 137 FPEALRGCEQESNIAFLIDGSGSINTIDFQKKEFVSTVMDQFQSKTLFSLMQYSDEF 196  
 QY 181 RIHFTFKFQNNPNPSLKIPTQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
 Db 197 RTHFTENDPKRNPDPKSHVRPQLNGRTKTASGIRKVVRELQKINGARDNAAKILVVI 256  
 QY 241 TDEKEKEDPLGYEDNVTPEADREGVIRYIVGVGDADFSEKSRQELNVTASKPPRDHVFQIN 300  
 Db 257 TDEKEKEDPLGYEDNVTPEADREGVIRYIVGVGDADFSEKSRQELNVTASKPPRDHVFQIN 316  
 QY 301 NFEALKTIONLRKEKIFALEGTGSSSFEHMSQEGFSAATSGNPLISTVGSVDWAG 360  
 Db 317 NFEALNTIIRNQEKFIFALEGTGSSSFEHMSQEGFSAATSGNPLISTVGSVDWAG 376  
 QY 361 GVFLYTSKEKSTFTNMTVRDSDMDNDAYLGAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420  
 Db 377 GAFLYPSKDKKAFINTTRDSDMDNDAYLGASAVISNRVQSLVLGAPRYQHIGLVAMFX 436  
 QY 421 QNTGMWESNANVKTQTGAIFGASLCSVDVDSNGSTDLVLGAPHYETQRTGGQVSCPL 480  
 Db 437 QNFGAWPEHTDKGSGISYFGASLCSVDMDADGNTNLILGAPHYETKRTGGQVSCPL 496  
 QY 481 PRGORARWQCDVLYGQGOQWGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
 Db 497 PRG-RARWQCEALHGDQHPGRFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAVYIF 555  
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGAGQHVLLRSQ 600  
 Db 556 HGASVASISTPHSQRIAGARFSPGLQYFGQSLGGKDLTMDGLMDLAVGAGQHLLLRQA 615  
 QY 601 PVLRVKAIMENPREVARNVFCNDQVVKGEAGEVRCVLHVQKSTRDRLRREGIOQSVVT 660  
 Db 616 PVLRLIATMBFSPKKVARSVFACQVLRKNDAGEVRCVLRVKNTRDRLRREGDIQSTVT 675  
 QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
 Db 676 YDLALDPGRSVVRAPFDETKNGILRRIRVFGTQKCETLKLIPDCVMSVSPITLRNY 735  
 QY 721 SLVGTPLSAFNLRLPVLAEADAORLTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
 Db 736 TLVGEPLRSRDLRLPVLAMEAQRIITAMFPFEKNCNDNICQDDLSITVSTRVNLVVG 795

```
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKYSTLONORSORSWRLACESASSTEV 840
Db 796 DSRDFDVSFVLRNDGEDSYTKVTCYYPVSGLSVRKVSASQNFQSKPWRVIAE-PPSSSEG 854
QY 841 SGALKSTSCSINHPFIPENSEVFNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900
Db 855 QGVLSKTIWINDHPFIPANSEVFNITFDVDSASLGKLLKANVTSENNVSRDTKTES 914
QY 901 QLELPVKYAVVWVTVSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSIPISLVFLVPV 960
Db 915 QLELPVKYAVVWVTVSGESSKYLNFTASENTSKVIHQYQFNVLQORSFPVSVVFWIPV 974
QY 961 RLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVUVNCSIAVCORIQCIP 1020
Db 975 QINKVTIWDPPQVTFSONLSVCTEQKSPSHSPQDELERTPVINCVAVCRIQCIP 1034
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSTLLPGQAFVRSQTE 1080
Db 1035 SFNSKEIFNVTLQGNLLFDWYIETSHDHLIVSTAEILFNDVSAFALLPGQETFKVKAQTE 1094
QY 1081 KVPEFVNPDLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGPPGAEPQ 1137
Db 1095 KVPEYTVHNPVLTIVGSSVGLVLLALITAGLYKLGFFKQYKDMNEAGQDGP 1151
```

## RESULT 4

```
Q28984 PRELIMINARY; PRT; 920 AA.
ID Q28984
AC Q28984;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CD11b (Fragment).
GN Name=CD11b;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
SEQUENCE FROM N.A.
RA Lee J.-K., Schook L.B., Rutherford M.S.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; U40072; AAB16869.1; -.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR00413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3-.
DR Pfam; PF00092; VWA; 1-.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00327; VWA_1.
DR PROSITE; PS02334; VWFA; 1.
KW Cell adhesion; integrin; Transmembrane.
FT NON_TER 1
FT TER 920
SQ SEQUENCE 920 AA; 102440 MW; E96CC51E350DD5AC CRC64;
```

```
Query Match 64.8%; Score 3809.5; DB 2; Length 920;
Best Local Similarity 78.9%; Pred. No. 3.2e-243;
Matches 727; Conservative 84; Mismatches 109; Indels 1; Gaps 1;
QY 118 PQKPEALRGCPQEDSIAFLVDGSGSIIPHDFRAREKFISTVMBQLKSKTLFSLMOYS 177
Db 1 PQKPEALRGCPQESDIAFLIDGSGSINRLDFQRMKEFVSTVMQGFQKSKTLFALMOYS 60
QY 178 EEFRIHFTPEFQNNPNPSLIKPTOLLGRTHATGIRKVVVRELFNITNGARKNAFKIL 237
Db 61 EDFYHFTFNDKRNPSPKLLVRPIQLLGRTHATGIRKVVVRELFSKSGSARENALKIL 120
```

```
QY 238 ILITDCEKFGDPLGYEDVIEPADREGVIRYVIGVDAFRSEKGRQELNTVASKPPRRDHV 297
Db 121 VVITDSEKFGDPLGYEDVIEPADRKGVIRYVIGVDAFNWSKREELNTIASKPCGDHVF 180
QY 298 QINNFEALKTIONQLEKIFAIEGTOTGSSSSSEHEMSQEGFSAATNSGFLLSITVGSYD 357
Db 181 QVNNFEAVKTIQOLQEKTEFAIEGTOTGSSSSSEHEMSQEGFSAATNSGFLLSITVGSYD 240
QY 358 WAGGVFLYTSKESKSTFINMTRVDSMDNDVLYGAAAIILNRYQSLVGLAPRYOHIGLVA 417
Db 241 WAGGAFULHMPKDRVIFINTTRVDSMDNDVLYGAAVEILNRQASLVGLAPRYOHTGLV 300
QY 418 MFRQNTGMBSNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQVSV 477
Db 301 MFKQNSGAEKNADIKGSQIGSYFGASLCSVDVNRDGSDDLVLIGAPHYEYQTRGGQVSV 360
QY 478 CPLPRGARAWQCDVLYGEGQGFWRGFGAALTVLGVNVDGKLTVDVAIGAPBEDNRGAV 537
Db 361 CPLPQG-RAKWQCRVILCGEQHPWRFGAALTALGDVNGDKLTVDVAIGAPGEOQNRGAV 419
QY 538 YLFHGTSGSGISPSHSORIASGKLSRLOVFGOSLGGQDLTMDGLVDLTVGAGHVL 597
Db 420 YLFHGTSELGSPSHSORIASGQSLRLOVFGOSLGGQDLTMDGLMDLAVGAGHVF 479
QY 598 RSQVLRKVKAIMBENPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQS 657
Db 480 RSQVLRVEASVMFKPREVARNVVECHQQAQTKIAGEVQVCLQVRKSTMDRLREGDQTS 539
QY 658 VVTVYDLALDSGRPHSRVAFNETKNSRRTOVGLGTOTCETLKLQLPNCIEDPVSPVLR 717
Db 540 IITYDLALDPGRPHPRVAFNETKNSRRTOVGLGTOTCETLKLQLPNCIEDPVSPVLR 599
QY 718 LNFSLVGTPLSAGNLPVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCL 777
Db 600 LNFSLVQKPASSGNLRLPVLAVDAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCL 659
QY 778 VVGSPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKYSTLONORSORSWRLACESAS 837
Db 660 VVGSPRDLKVTLTVRNGEDSYRTQVTFPPDLVSRKYSTLONORSORSWRLACESAS 719
QY 838 TEVSGALKSTSCSINHPFIPENSEVFNITFDVDSKASLGKLLKANVTSENNMPRTN 897
Db 720 TEESTALKSTSCSINHPFIPDNSEVFNITFDVDPADFLYKLLKANVTSENNMPSSNK 779
QY 898 TEFQLELPVKYAVVWVTVSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSIPISLV 957
Db 780 TEFQLELPVKYAVVWVTVSLEVSTKYFNFTASEKTRHVEHQYQFNVLQORSIPISLV 839
QY 958 VPVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVUVNCSIAVCORI 1017
Db 840 VPVRLNQTIVWDRPQVTFSENLSTCHTKERLPSHSDFLAELKAPVUVNCSIAVCORI 899
QY 1018 DIPFGIOEEFNATLKGNSL 1038
Db 900 DIPSGIOEEFNATLKGNSL 920
```

## RESULT 5

```
ITAX HUMAN STANDARD; PRT; 1163 AA.
ID ITAX HUMAN
AC P20702; Q8IVA6;
DT 01-PEB-1991 (Rel. 17, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Lue M5).
GN Name=ITCAX; Synonyms=CD11c;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
SEQUENCE FROM N.A.
```

RX MEDLINE=88166645; PubMed=3327687;  
 RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;  
 RT "CDNA cloning and complete primary structure of the alpha subunit of a  
 RL leukocyte adhesion glycoprotein, p150,95.";  
 RN ENBO J. 6:4023-4028 (1987).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90153906; PubMed=2303426;  
 RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;  
 RT "Genomic structure of an integrin alpha subunit, the leukocyte p150,95  
 RL molecule.";  
 RN J. Biol. Chem. 265:2782-2788 (1990).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Tothiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Rosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gumaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski S.I., Skalska U., Smailus D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RL and mouse cDNA sequences.";  
 RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 [5]  
 RP SEQUENCE OF 20-43.  
 RX MEDLINE=87167596; PubMed=3549901;  
 RA Miller L.J., Wiebe M., Springer T.A.;  
 RT "Purification and alpha subunit N-terminal sequences of human Mac-1  
 RL and p150,95 leukocyte adhesion proteins.";  
 RN J. Immunol. 138:2381-2383 (1987).  
 CC -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It  
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell  
 CC interaction during inflammatory responses. It is especially  
 CC important in monocyte adhesion and chemotaxis.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X  
 CC associates with beta-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and  
 CC granulocytes.  
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M81695; AAA51620.1; -.  
 DR EMBL; M29165; -; NOT ANNOTATED CDS.  
 DR EMBL; M29487; AAA51620.1; AUT\_SEQ.  
 DR EMBL; M29482; AAA51620.1; JOINED.  
 DR EMBL; M29483; AAA51620.1; JOINED.  
 DR EMBL; M29484; AAA51620.1; JOINED.  
 DR EMBL; M29485; AAA51620.1; JOINED.  
 DR EMBL; M29486; AAA51620.1; JOINED.  
 DR EMBL; BC038237; AAH38237.1; -.  
 DR FIR; A36584; RWHUIC.  
 DR PDB; 1N3Y; X-ray; A=141-338.  
 DR Genew; HGNC:6152; ITGAX.  
 DR MIM; 151510; -.  
 DR GO; GO:0008305; C:integrin complex; TAS.  
 DR GO; GO:0004872; F:receptor activity; TAS.  
 DR GO; GO:0007155; P:cell adhesion; TAS.  
 DR GO; GO:0009887; P:organogenesis; TAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF01839; FG-GAP\_3.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWFA; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWFA\_1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS0234; VWFA; 1.  
 KW 3D-structure; Calcium; Cell adhesion; Direct protein sequencing;  
 KW Glycoprotein; Integrin; Magnesium; Polymorphism; Receptor; Repeat;  
 KW Signal; Transmembrane.  
 FT SIGNAL 1 19  
 FT CHAIN 20 1163 Integrin alpha-X.  
 FT DOMAIN 20 1107 Extracellular (Potential).  
 FT TRANSMEM 1108 1128 Potential.  
 FT DOMAIN 1129 1163 Cytoplasmic (Potential).  
 FT REPEAT 34 87 FG-GAP 1.  
 FT REPEAT 88 ? FG-GAP 2.  
 FT DOMAIN 165 351 VWFA.  
 FT REPEAT ? 401 FG-GAP 3.  
 FT REPEAT 402 453 FG-GAP 4.  
 FT REPEAT 455 517 FG-GAP 5.  
 FT REPEAT 518 576 FG-GAP 6.  
 FT REPEAT 581 633 FG-GAP 7.  
 FT CA\_BIND 466 474 Potential.  
 FT CA\_BIND 530 538 Potential.  
 FT CA\_BIND 593 601 Potential.  
 FT SITE 1131 1135 GFFKR motif.  
 FT DISULFID 69 76 By similarity.  
 FT DISULFID 108 126 By similarity.  
 FT DISULFID 655 712 By similarity.  
 FT DISULFID 771 777 By similarity.  
 FT DISULFID 848 863 By similarity.  
 FT DISULFID 998 1022 By similarity.  
 FT DISULFID 1027 1032 By similarity.  
 FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 392 392 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 697 697 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 735 735 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 899 899 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 939 939 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 1050 1050 N-linked (GlcNAc...) (Potential).  
 FT VARIANT 48 48 W -> R (in dbSNP:11574633).  
 FT /FTID=VAR\_018672.  
 FT CONFLICT 209 209 T -> S (in Ref. 4).  
 FT CONFLICT 251 251 T -> A (in Ref. 4).  
 FT CONFLICT 469 469 T -> S (in Ref. 4).  
 FT CONFLICT 490 490 G -> A (in Ref. 2).  
 FT CONFLICT 547 547 E -> K (in Ref. 4).  
 FT CONFLICT 756 756 D -> L (in Ref. 1).  
 FT CONFLICT 819 819 I -> V (in Ref. 4).  
 FT CONFLICT 1161 1163 SEK -> TPHYPQDNV (in Ref. 4).  
 FT STRAND 150 150

```
FT TURN 160 161
FT HELIX 164 178
FT TURN 179 180
FT TURN 183 185
FT STRAND 186 193
FT STRAND 197 201
FT HELIX 203 208
FT HELIX 212 216
FT TURN 217 218
FT STRAND 226 226
FT HELIX 228 236
FT TURN 237 240
FT HELIX 242 244
FT TURN 245 245
FT TURN 248 249
FT STRAND 251 258
FT STRAND 263 263
FT HELIX 269 278
FT TURN 279 280
FT STRAND 282 288
FT HELIX 290 293
FT TURN 296 297
FT HELIX 298 304
FT HELIX 310 312
FT STRAND 313 316
FT HELIX 319 325
FT HELIX 326 334
FT TURN 335 335
SQ SEQUENCE 1163 AA; 127886 MW; 83658A13B5C5DE8F CRC64;

Query Match 59.2%; Score 3479; DB 1; Length 1163;
Best Local Similarity 61.1%; Pred. No. 3.6e-221;
Matches 690; Conservative 141; Mismatches 292; Indels 6; Gaps 4;

Qy 1 ENLDTENAMTFOENARGFGSVWQLOGSRVVVVGAPQIIVAAANQRGSLYQCDYSTGSCPEI 60
Db 20 ENLDTEELTAFRVDSAGFGDSVVQYANSVVVVGAPQIKTAANQTGGLYQCGYSTGACEPI 79

Qy 61 RLQVPVAVNMVSLGSLAATSPOLLACGTVHQTCSNTYVVGKLCFLFGSNLRQPOK 120
Db 80 GLQVPPVAVNMVSLGSLASTTSPSOLLACGTVHHECGRNMYLTGLCLFLGPT--QLTOR 137

Qy 121 FPEALRGCPQSDIAFIADVDSGSIIPHDPRAKEFTSTVMEQLKSKTKLPSLMQYSEEF 180
Db 138 LPVSRQCPQEQDIFVLIDGSGSISSRNFAFMNFRAVISQFQSTQPSLMQFNKF 197

Qy 181 RIHFTFKFQNNPNSRIKPIITQLLGRTHATGIRKVVVRELFNITNGARKNAFKILILI 240
Db 198 QTHFTFEFRRTSNPLSLLASVHQLOGFTYTATAIQNVVHRLFHASYGARRDATKILIVI 257

Qy 241 TDGKFGDPLGYEDVIPEADREGVIRVIGVGDAPRSEKSRQELNTVASKPRDRHVQIN 300
Db 258 TDGKFGDGLDYKDVIPMADAAGIIRYAIGVGLAFQNRNSWKELNDIASKPSQEHIFKE 317

Qy 301 NFEALKTTQNLREKIPIAETGTGSSSFFHEMSQEGFSAATISNGPLLTSTVGSYDWAG 360
Db 318 DFDALKDITQNLREKIPIAETGTGSSSFFHEMSQEGFSAATISNGPLLTSTVGSYDWAG 377

Qy 361 GVFLYTSKEKSTFTINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLAPRYQHILVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVDMDSYLGSTELALWKGVSQSLVGLAPRYQHTKAVIFT 437

Qy 421 QNTGHWESNANVKGTQIGAYFGASLCSVDVDSNSTDLVLIGAPHYEOTRGQSVSCLP 480
Db 438 QVSRQWRKAKAVTGTQIGSYFGASLCSVDVDTGSDTLVLIGAPHYEOTRGQSVSCLP 497

Qy 481 PRGQARWQCDVLYGEGQGWGRFGAALTIVLGVNGDKLTDVAIGAPGEEDNMGAVYLF 540
Db 498 PRGWR-RWNCDAVLYGEGQGWGRFGAALTIVLGVNGDKLTDVIGAPGEENRGAVYLF 556

Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFQSLSGGQDLTMGLVDLTVGAQGHVLLRSQ 600
Db 557 HGVLTGFSISPSHSQRIAGSKLSPRLQYFQSLSGGQDLTMGLVDLTVGAQGHVLLRSQ 616
```

```
Qy 601 PVLRVKAIMFEFNPREVARNVFECNDQVWKGKAGEVVRVCLHVOKSTEDRLREGIQSVVT 660
Db 617 PVLWVGVSQFIPAEIPRSFAFECREQVVSQETLVQSNICLYIDKRKNLLGSRDLQSSVT 676

Qy 661 YDLALDSGRPHSRVAFNETKNSTRRTQVTLGLTQTCTETLKLQLENCIEDPVSPIVRLNF 720
Db 677 LDIALDGRSLSPRATQETKNRSLSRVRLGLKAHCENFNLLPSCVEDSVTPITRLNF 736

Qy 721 SLVGTPLSAPGNLRPVLAEDAORLFTALFFEFKNCNDNTICODDLSITTFMSLDCLVVG 780
Db 737 TLVGKPLLAFLNLRPMLAADAQRYFTASLFFEFKNCADHTICODNLGISFSPGLKSLVG 796

Qy 781 GPREFVTVTRNDGDSVYRTQTFEFPPLDLSYRKVSTLQNRSQSRMSRLACSSASSTEV 840
Db 797 SNLELNAEVMVNDGEDSYGTTITFHPAGLSYRYVABGQKQGLRLHLHITCDAPVGV-- 854

Qy 841 SGALKSTSCSINHPIFPENSEVITFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db 855 SQGTWSTSCRINHILIFRGGAQITFLATFDVSPKAVLGDRLLTANVSSENNTPTSKTTF 914

Qy 901 QLELPVKYAVYVMVTSKGVSTKYNLTAS-ENTSRVMQHOYQYNSLQORSIPISLVFLVP 959
Db 915 QLELPVKYAVYTVVSSHEQFTKYNLFSESEKESHVAMHRYQVNNLQORLDPVSNFVWP 974

Qy 960 VRLNQTVIWDPRPOVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDI 1019
Db 975 VELNQEAVMVDVEVSHFQNFSLRCSSEKIAPPASDFLAHKQNPVLDCSAGCLFRCDV 1034

Qy 1020 PFGIQIEFNATLKNLSFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPGQAFVRSQTE 1079
Db 1035 PSFSVQELDFTLKNLSFGWVRQILQKQVSVSVASITFDTSVYSLPGQEAFFMRAQTT 1094

Qy 1080 TKVEPFPVPPLPLIVGSSVGGLLLLALITAAALYKLGFFRQYKDMWE 1128
Db 1095 TVLEKYKVNPTPLIVGSSIGLLLLALITAVLYKGVFFRQYKEMEE 1143

RESULT 6
ITAD HUMAN
ID ITAD HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor (Leukocyte integrin alpha D) (CD11d) (ADB2).
GN Name=ITGAD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=36111956; PubMed=8777714;
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukocyte integrin, alpha d beta 2, binds preferentially to ICAM-3."
RL Immunity 3:683-690 (1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin gene CD11d. Essential role of Sple and Sp3."
RN J. Biol. Chem. 275:8959-8969 (2000).
RN [3]
RP SEQUENCE OF 467-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., Lebeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human beta 2-integrin alpha subunit."
RT
```

Gene 171:291-294(1996).  
[4]  
RN INTERACTION WITH VCAM1.  
RP MEDLINE=99059842; PubMed=9841932;  
RX Grayson M.H., Van der Vieren M., Sterbinsky S.A., Michael Gallatin W.,  
RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
EA "alpha2beta2 integrin is expressed on human eosinophils and functions  
RT as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-  
RT 1).";  
RL J. Exp. Med. 188:2187-2191(1998).  
[5]  
RN INTERACTION WITH VCAM1.  
RP MEDLINE=99370002; PubMed=10438935;  
RX Van der Vieren M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
EA "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
RT binding interface between I domain and VCAM-1.";  
RL J. Immunol. 163:1984-1990(1999).  
CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and  
CC VCAM1. May play a role in the atherosclerotic process such as  
CC clearing lipoproteins from plaques and in phagocytosis of blood-  
CC borne pathogens, particulate matter, and senescent erythrocytes  
CC from the blood.  
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D  
CC associates with beta-2.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell  
CC lines and subsets of peripheral blood leukocytes and strongly on  
CC tissue-specialized cells, including macrophages foam cells within  
CC atherosclerotic plaques, and on splenic red pulp macrophages.  
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
CC with I-domains do not undergo protease cleavage.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial/  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U37028; AAB38547.1; -;  
DR EMBL; U40274; AAB60634.1; -;  
DR EMBL; U40275; AAB60635.1; -;  
DR EMBL; U40276; AAB60636.1; -;  
DR EMBL; U40277; AAB60637.1; -;  
DR EMBL; U40279; AAB60638.1; -;  
DR EMBL; U40278; AAB60638.1; JOINED.  
DR EMBL; AF187881; AAF62875.1; -;  
DR HSSP; P11215; IHHQ  
DR Genew; HGNC:6146; ITGAD.  
DR MIM; 602453; -;  
DR GO; GO:0008305; C:integrin complex; TAS.  
DR GO; GO:0016337; P:cell-cell adhesion; NAS.  
DR GO; GO:0007160; P:cell-matrix adhesion; NAS.  
DR GO; GO:0006955; P:immune response; NAS.  
DR InterPro; IPR000413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_alpha; 1.  
DR Pfam; PF00092; VWFA; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWFA; 1.  
DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
DR PROSITE; PSS0234; VWFA; 1.  
KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
KW Repeat; Signal; Transmembrane.  
FT SIGNAL 1 17  
FT Potential.

```
QY 481 PRGORARWOCDAVLYCEGQCPWGRFGAALTVLGDVNGDKLTDTVAIGAPGEDNRGAVYLF 540
Db 497 PRGORVOWQCDADVLRGEQGHWPGRFGAALTVLGDVNGDKLTDTVAIGAPGEDNRGAVYLF 556
QY 541 HGTSGSGISPSHSORISAGSKLSRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGASESGISPSHSORISAGSKLSRLQYFGQALSGGQDLTMDGLMDLAVGARGQVLLRSL 616
QY 601 PVLVRVKAIMEFNPREVARNVFECDQVVKQKEAGEVRVCLHVQKSTRDRRLRGEIQSVVT 660
Db 617 PVLKVGVMARFSEVEKAVYRCWEKPKSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 674
QY 661 YDLALDSGRPHSRVAVNETKNSRTRQTVGLTCTCETLKLQLPNCIEDPVSPIVLRINF 720
Db 675 FDUALDPGRUTSRAINFNETKNTLTKRRTKLTGLGHCETLKLKLLPDCVEDVSPILHLNF 734
QY 721 SLVGTPLSAGFNLRLPVLAEQAORLFTALPFKNCNGNDNICQDDLSTIFSMSLDCLVVG 780
Db 735 SLVREPIPSQNLRLPVLAVGSQDLFTASLPFEKNCQDGLCEGDLGVTLSFGQLTLTVG 794
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLSTYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db 795 SSLELNIVTVWVWAGDSYGVVSLYYPAGLSHRVSGAQKPHQSAIRLACETV-PTED 853
QY 841 SGALKSTCSINHPPIPEENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTKNTPE 900
Db 854 EG-LRSRCVSNHPPIHEGNGTFIVTFDVSYKATLGDRLMLKASSENKAKSSKATF 912
QY 901 QLELPVKYAVYVMTSHGVSTKYNLF-TASENTSRVMQHOYQVNSLQSRSLPISLFLVP 959
Db 913 QLELPVKYAVYVMTSROESTKYFNFATSDKKWKAEBHYRVNNLSQRLAITSINFWP 972
QY 960 VRLNQTVINDRPQVTPSENLSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCORQCDDI 1019
Db 973 VLLNGVAVWVDMVMEAPSQSL--PCVSRKPPQHSDFLTQISRSPMLDCSIADCLQFRCDV 1030
QY 1020 PFGIGIEEFNATLKGNSLSPDWYIKTSHNLLIYSTAEILFENDSVFTLLPGQAFVRSOTE 1079
Db 1031 PSFSVQEEFLTUKGNLSFQWVRETQKKVLLVSVVAEITPDTYSQLPQGEAFRAQME 1090
QY 1080 TKVEPEFVNPLFLIVGSSVGGLLLLALITAAALYKLGFFRQYKQKDMNSE 1128
Db 1091 MVLEEDENVYNAIPIIMGSSVGALLLALITATLYKLGFFRHYKEMLED 1139
QY 1091 MVLEEDENVYNAIPIIMGSSVGALLLALITATLYKLGFFRHYKEMLED 1139
Db 1091 MVLEEDENVYNAIPIIMGSSVGALLLALITATLYKLGFFRHYKEMLED 1139
RESULT 7
Q6KAS4
ID Q6KAS4 PRELIMINARY; PRT; 1188 AA.
AC Q6KAS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (Fragment).
GN Names=MFLJ00114;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kituno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Suga Y., Kitamaga H., Nakagawa T., Nagase T., Ohara O., Koga H.,
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL DNA Res. 11:167-180(2004).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL, AK131133; BAD21383.1; -.
DR GO; GO:0009897; C:external side of plasma membrane; IDA.
DR InterPro; IPR000413; Integrin_alpha.
```

```
DR InterPro; IPR0002035; VWF A.
DR Pfam; PF01839; FG-GAP; 3_
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR04453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
DR PROSITE; PS0234; VWA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON_TER 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDB4178 CRC64;

Query Match 56.7%; Score 3329.5; DB 2; Length 1188;
Best Local Similarity 56.7%; Pred. No. 3e-211;
Matches 649; Conservative 170; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQSGSRVVVGAPOEIVAAANQSGSLYQCDYSTGCEPI 60
Db 39 FNLDSEKPTFHMDDGAEFGHSLVLYDSSVWVVGAPKEIKATNIGGLYKGYHTGCEPI 98
QY 61 RLQVPVEAVNMSLGLSLAATTGPPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPK 120
Db 99 SLQVPPEAVNMSLGLSLAATNPSSLACGPTVHHTCRENIYLTGLCFLLSLSSFKQS-QN 157
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHFERRAKEFISTVMEQLKSKTLFSLMOYSEEF 180
Db 158 FPTAQOCECPQDQIVFLIDGSGSISTDFEKMFLDFKAVMSQLQRPSTFSLMQFSDYF 217
QY 181 RHFTPEKPEQNNPRSLIKPTQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 218 RVHFTNNFISTSPSLDLSVRQLAGTYTASAIKHVITELFTTSGARQADATKVLIV 277
QY 241 TDGEKFGDPLGVEDVIPEADREGVIRYIVGVGDAPFSEKSRQBELNTVASKPPRDHVFIQ 300
Db 278 TDGRKQGDNLSDYSPVMAEASIIRYAIGVGKAFVNEHSKQBLKAIASMPSEHYVFSVE 337
QY 301 NPEALKTIQNLBEKIFAETGTGSSSEFHEMSQEGFSAATSTNGPLLSITVGSYDWAG 360
Db 338 NFDALDKIENQLKEKIFAETGTTPSSSTFELEMSQEGFSAVFTPGVPVLGAVGSEFSWSG 397
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAIILNRNVQSLVGLAPRYQHTGLVAMPR 420
Db 398 GAFLYFSNWRPTFINNSQENEDMDAYLGYSTALAPWKGVHSLILGAPRHQHTKGVIFT 457
QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGQGVSVCP 480
Db 458 QESRHRPRKSEVRGTQIGSYFGASLCSVDMDRGSTDLVLIGVPHYEHTRGQGVSVCPM 517
QY 481 PRGORARWOCDAVLYCEGQCPWGRFGAALTVLGDVNGDKLTDTVAIGAPGEDNRGAVYLF 540
Db 518 P-CVGRRWCHGTTLHGEQHPWGRFGAALTVLGDVNGDSLADVAIGAPGEENRGAVYIF 576
QY 541 HGTSGSGISPSHSORISAGSKLSRLQYFGQSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 577 HGASRODIAPSPQISASQIPSKIQYFGQSLSGGQDLTMDGLVDLAVGSKGVLLLR 636
QY 601 PVLVRVKAIMEFNPREVARNVFECDQVVKQKEAGEVRVCLHVQKSTRDRRLRGEIQSVVT 660
Db 637 PILRVSPVTHFTPAETISRSVFECQVAPETLSDATVCLIHESPKTQL--GDLRSTVT 694
QY 661 YDLALDSGRPHSRVAVNETKNSRTRQTVGLTCTCETLKLQLPNCIEDPVSPIVLRINF 720
Db 695 FDUALDHGRULSTRAIFKETKTRALTRVKTLGLNKHCSVKLLLPACVDSVTITLRLNF 754
QY 721 SLVGTPLSAGFNLRLPVLAEQAORLFTALPFKNCNGNDNICQDDLSTIFSMSLDCLVVG 780
Db 755 SLVGVPEISSLQNLQPLAVDDQTYFTASLPFEKNCAGADHICQDDLSSVVFPGDPLKTLVVG 814
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLSTYRKVSTLQNRQSRQSWRLACESASSTEV 829
Db 815 SDLELNVDVTVSRNDGEDSYGTTVTLFPVGLSFRVAVGQVFLRKKEDDQWQRKRGQSLH 874
```

```

QY 830 LACESASSTVSGALKSTSCSINHPIFPENSEVENTITDVSOKASIGNKLKLLKANVTSE 889
Db 875 LMCD--STDRSQGLWSTSCSRHVIFRGSGMTFLVFDVSPKAEIGDRLLLRARVGE 932
QY 890 NNMPRTNKTFFOLELPVKYAVYVMTVSHGVSTKYNFTASE-NTSRVMOHQYQVSNLQOR 948
Db 933 NNVPGETKTTFFOLELPVKYAVYVMTVSHGVSTKYNFTASE-NTSRVMOHQYQVSNLQOR 992
QY 949 SLFISLVFLVPLRLNQTIVWDRPOVTFSENLSTCHTKERLPKSHDPLAELKAPVNVCS 1008
Db 993 DVPVSINFVWPIELKGEAVW-TVMVSHPOPLTQCYRNRLKPTQDFDLTHMQSPVLDCS 1051
QY 1009 IAVCQRIQCDIPFPGIQEENATLKNGLSPDWIKTSHNHLIVSTAELFNDVSFTLLP 1068
Db 1052 IADCLHLRCDIPSLGILDELFLYFKGNLSFGWISQTLQKKVLLSBAITFTNTSVYSQLP 1111
QY 1069 GQAFVRSQTEKTVKPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMWSE 1128
Db 1112 GQAFVRSQTEKTVKPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMWSE 1171

RESULT 8
ID BAD21383
AC BAD21383;
DT 01-JUN-2004 (T-EMBLrel. 27, Created)
DT 01-JUN-2004 (T-EMBLrel. 27, Last sequence update)
DE MFLJ00114 protein (Fragment).
GN MFLJ00114.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kitano R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs
RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries."
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK131133; BAD21383.1;
FT NON_TER
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match
Best Local Similarity 56.9%; Pred. No. 3e-211;
Matches 649; Conservative 170; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFQENARGFGOSVVLQGSRRVVGAPQEIIVAAORGSLYQCDYSTGSCPTI 60
Db 39 FNLDAEKPTFHMDGAEFGHVSFLQYDSSVWVVGAPKEIKATNQIGGLYKCGYHTGNCPTI 98
QY 61 RLQVPEAVNMISGLSLAATTSPQILACGPTVHTQCSNTYVKGICFLFGSNLRQOPQK 120
Db 99 SLQVPEAVNMISGLSLAATNPSWILLACGPTVHTQCSNTYVKGICFLFGSNLRQOPQK 157
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEFTISTVMEQLKSKTLFSLMQYSEEF 180
Db 158 FPTAQECQKQDQDIIVFLIDGSGSISTDEKMLDFKAVMSQIQRPSTFSLMQYSDYF 217
QY 181 RHFTTKEFQNNPNSLKIPIQLGRHTATGIRKRVRELFTNNGARKNAFKILILI 240
Db 218 RVHFTFNNFTSPLSLSDSVRLRGYTYTASAIKHVITELFTTQSGARQDQATKVLIV 277
QY 241 TDGEKGDPLGYEDVLEADREGVIRVVGDAFSEKSRQELNTVASKPPRDHVFQIN 300
Db 278 TDGRKQGDNLSDYSVDPMAEASIIIRYAGVGKAFYNESKQELKAIASMPSEHYVFSVE 337
QY 301 NFALKTIQNLREKIFAIEGTTPSSSTPELEMSQEGFSAVTTPDGPVLGAVGSPWSG 397

```

```

Db 338 NFALKTIQNLREKIFAIEGTTPSSSTPELEMSQEGFSAVTTPDGPVLGAVGSPWSG 397
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAIILNRRVQSLVGLAPRYOHIGLVAMFR 420
Db 398 GAFLYPSNRPTFINNSQENEDMDAYLGYSTALAFWKGVHSLILGAPRHQTKGVIFT 457
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLGAPHYYEQTGRGQVSCPL 480
Db 458 QESRHRWPKESEVGTQIGSYFGASLCSVDMDRDGSTDVLVLGAPHYYEHTRGQVSCPM 517
QY 481 PRQARWOCDAVLYGEGQGPWGRFGAALTVLGDVANGDKLTDVAIGAPGEDNRGAYVLF 540
Db 518 P-GVGRWHCCITLHGEQGHWPGRFGAALTVLGDVANGDSLADVAIGAPGEENRGAYVIF 576
QY 541 HGTSGSISPSHSQRIAGSKLSRLQYFGQSLSGGGDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 577 HGASRQDIAPSPQISASQIPSKIQYFGQSLSGGGDLTMDGLVDLTVGAQGHVLLRSQ 636
QY 601 PVLARKAIMEFNPVARNVFECDQVVGKAGEVRVCLHVOKSTRDLRREGQIOSVVT 660
Db 637 PILRVSPVHTFPAEISRSVFECQVAPQTLSDATVCLHSHESPKTLQ--GDLRSTVT 694
QY 661 YDLALDSGRPHRAVENETKNSRROTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720
Db 695 FDLALDHGRLSTRAIFKETKTRALTVKTLGLNKHCHESVLLLPACVEDSVPTITRLNF 754
QY 721 SLVGTPLSAFNGLRPVLAEADQRLFTALPFFPKKNCNDNICODDLSITFSMSLDCLVVG 780
Db 755 SLVGVPISSILQNLQPLAVDDQTYFTASLPFEKNCADHICODDLSVWFGFPLDKTLVVG 814
QY 781 GPREFNVTVRNDEGDSYRTQVTFPFDLDSVRKYSTLQ-----NORSQSWR 829
Db 815 SDLELVNDVTNSDGEDSYTTLVTFYPVGLSPRRVARGQVFLRKXEDQOWQRGQSHLH 874
QY 830 LACESASSTVSGALKSTSCSINHPIFPENSEVENTITDVSOKASIGNKLKLLKANVTSE 889
Db 875 LMCD--STDRSQGLWSTSCSRHVIFRGSGMTFLVFDVSPKAEIGDRLLLRARVGE 932
QY 890 NNMPRTNKTFFOLELPVKYAVYVMTVSHGVSTKYNFTASE-NTSRVMOHQYQVSNLQOR 948
Db 933 NNVPGETKTTFFOLELPVKYAVYVMTVSHGVSTKYNFTASE-NTSRVMOHQYQVSNLQOR 992
QY 949 SLFISLVFLVPLRLNQTIVWDRPOVTFSENLSTCHTKERLPKSHDPLAELKAPVNVCS 1008
Db 993 DVPVSINFVWPIELKGEAVW-TVMVSHPOPLTQCYRNRLKPTQDFDLTHMQSPVLDCS 1051
QY 1009 IAVCQRIQCDIPFPGIQEENATLKNGLSPDWIKTSHNHLIVSTAELFNDVSFTLLP 1068
Db 1052 IADCLHLRCDIPSLGILDELFLYFKGNLSFGWISQTLQKKVLLSBAITFTNTSVYSQLP 1111
QY 1069 GQAFVRSQTEKTVKPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMWSE 1128
Db 1112 GQAFVRSQTEKTVKPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMWSE 1171

RESULT 9
ID ITAX MOUSE
IT ITAX MOUSE STANDARD; PRT; 1169 AA.
AC Q9QX4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c).
GN Name=Itgax;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dendritic cell;

```



RA Huang X., Gorski K., Tong C., Rattis F.-M., Tseng S.-Y., Pardoll D.,  
 RA Tsuchiya H.;  
 RT "Isolation of genes selectively expressed by dendritic cells.";  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It  
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell  
 CC interaction during inflammatory responses. It is especially  
 CC important in monocyte adhesion and chemotaxis (By similarity).  
 CC -1- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X  
 CC associates with beta-2 (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -1- SIMILARITY: Contains 1 VWFA domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AF211864; AAF23492.1; --  
 CC HSP; P20702; IN3Y.  
 CC MGD; MGI:196509; Itgax.  
 CC GO; GO:0009897; C:external side of plasma membrane; IDA.  
 CC InterPro; IPR00413; Integrin\_alpha.  
 CC InterPro; IPR002035; VWF\_A.  
 CC Pfam; PF01839; FG-GAP; 3.  
 CC Pfam; PF00357; Integrin\_alpha; 1.  
 CC Pfam; PF00092; VWFA; 1.  
 CC PRINTS; PR01185; INTEGRINA.  
 CC PRINTS; PR00453; VWFADOMAIN.  
 CC SMART; SM00191; Int\_alpha; 5.  
 CC SMART; SM00327; VWFA; 1.  
 CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 CC PROSITE; PS50234; VWFA; 1.  
 CC Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 19 Potential.  
 FT CHAIN 20 1169 Integrin alpha-X.  
 FT DOMAIN 20 1116 Extracellular (Potential).  
 FT TRANSMEM 1117 1137 Potential.  
 FT DOMAIN 1138 1169 Cytoplasmic (Potential).  
 FT REPEAT 34 87 FG-GAP 1.  
 FT REPEAT 88 ? FG-GAP 2.  
 FT DOMAIN 152 330 VWFA.  
 FT REPEAT ? 402 FG-GAP 3.  
 FT REPEAT 403 454 FG-GAP 4.  
 FT REPEAT 456 518 FG-GAP 5.  
 FT REPEAT 519 577 FG-GAP 6.  
 FT REPEAT 582 634 FG-GAP 7.  
 FT CA\_BIND 467 475 Potential.  
 FT CA\_BIND 531 539 Potential.  
 FT CA\_BIND 594 602 Potential.  
 FT SITE 1140 1144 GFFKR motif.  
 FT DISULFID 69 76 By similarity.  
 FT DISULFID 108 126 By similarity.  
 FT DISULFID 656 711 By similarity.  
 FT DISULFID 770 776 By similarity.  
 FT DISULFID 858 873 By similarity.  
 FT DISULFID 1007 1031 By similarity.  
 FT DISULFID 1036 1041 By similarity.  
 FT CARBOHYD 89 89 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 267 267 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 393 393 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 734 734 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 949 949 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 1059 1059 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 1084 1084 N-linked (GlcNAc... ) (Potential).

SQ SEQUENCE 1169 AA; 129150 MW; C616412033C219A6 CRC64;  
 Query Match 56.58; Score 3320.5; DB 1; Length 1169;  
 Best Local Similarity 56.8%; Pred. No. 1.1e-210;  
 Matches 647; Conservative 172; Mismatches 302; Indels 19; Gaps 7;  
 QY 1 FNLDTENAMTFQBNARGFGQSVVQLQSGRVVVGAPQBIIVAAQORGSLSYQCDYSTGSCPEI 60  
 DB 20 FNLDAAEKLTHFMDGAFGHVLSVLYQDSSVWVVGAPKEIKATNQIGGLYKGYHTGNCPEI 79  
 QY 61 RLQVPVAVNMSLGLSLAATSPQLLIACGPTVHQTCTSENYVKGCLFCLFGSNLRQOPQX 120  
 DB 80 SLQVPPAVANISLGLSLAATNPSSLACGPTVHHTCRENIYLTGLCFLLSSSSFKQS-QN 138  
 QY 121 FPEALRCPCQEDSDIAFLVDGSGSIIPHDFERRAKERISTVWEOLKSKTKLPSLMQYSEEF 180  
 DB 139 FPTAQOQCPKQDDIVFLIDGSGSIISTDEKMLDFVKAVMQIQORSTSPSLMOPSDYF 198  
 QY 181 RIHFTFEKQNNPNRSLIKPIITQLLGRTHATGIRKVVRELFINITNGARKNAFKIILI 240  
 DB 199 RVHFTFNNFISTSSPLSLGSLVRLGYTYTASAIKHVITELFTTQSGARQADATKVLIVI 258  
 QY 241 TDGEKFGDPLGYEDVPEADREGVIRVIGVGAFAFRSEKSRQELNLTVASKPPDRHVFQIN 300  
 DB 259 TDRGRKQGDNLSDYVIPMAEAAAIIRYIAIGVKAFAFNEHKSQELKALIASMPHSHEYVSFE 318  
 QY 301 NFEALKTIONQLREKIFAIEGTOTGSSSSFEHMSQSGFSAAITNSGPLLSTVGSYDWAG 360  
 DB 319 NFDALKDIEQLKEKIFAIEGTETPSSSTFELEMSQSGFSAVFTPDGFVLGAVGSFWSG 378  
 QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGVAIAIIILNRRVQSLVGLGAPRYQHIGLVAMFR 420  
 DB 379 GAFLYPSNMRPTFINNSQENEDMRDAYLGYSTALAFWKGVHSLILGAPRHQHTKGVVFT 438  
 QY 421 QNTGMESNANVGTQIGAVFGASLCSVDVDSNGSTDLVIGAPHYETQRTGGQSVCLP 480  
 DB 439 QESRHWKPKGEVRGTQIGSYFGASLCSVDMDRDSGLDVLIGVPHYETRTGGQSVCPM 498  
 QY 481 PRGORARWQCDVILYGGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRAVYLF 540  
 DB 499 P-GVGSRWHCCTLHGEQHPWGRFGAALTVLGVDNGDSLADVAIGAPGEENRAVYIF 557  
 QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGHVLRLRSQ 600  
 DB 558 HGASRQDIAPSPQRISASQIPSRIOYFGQSLGGQDLTRDGLVDLAVGSKGRVLLRLTR 617  
 QY 601 PVLRVKAIMEFNPREVARNPFCNDQVVKKEAGEVRVCLHVQKSTDRDLREGIQSVWT 660  
 DB 618 PILRVSPVHTFAETSRVSFEQEQVAPEQTLSDATVCLHIESPKTL--GDLRSTVT 675  
 QY 661 YDLALDSGRPHSRAVENETKSTRTQVGLTQTCETLKLQLPNCIEDPVSPIVLRLNF 720  
 DB 676 FDIALDHGRSLSTRAIFKETKTRALTRVTKLGNKHCSVKLLLPACVEDSVPTILRLNF 735  
 QY 721 SLVGTPLSAFGNLRPVLAEDAQRLLFTALFPFEKXCGNDNICODDLSTITFSFMSDCLVVG 780  
 DB 736 SLVGVPISSLQNLQPLAVDDQTYETASLPEKXCGADHICQDDLSSVVFPGFDLKLTVVG 795  
 QY 781 GPREFNVTVVRNDGDSYRTQVTFPPDLSTRKVSTLQ-----NQBSQSWR 829  
 DB 796 SDLELNVDVTVSNDDGDSYGTITVTLFYFVGLSPRRVAEGQVFLKKEDQWQRRGQSLH 855  
 QY 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSE 889  
 DB 856 LMCD--STPDRSQGLWSTSCSRHVIIPRGSGQMTFLVTVSPKAEGLDRLLLRARVGE 913  
 QY 890 NNMPRTNKTEFOLELPVKYAVMVVTVSHGVSTKYLNFASE-NTSRVMQHQYQVSNLQGR 948  
 DB 914 NNVPGTPTKTFQLELPVKYAVYTMISSHDQFTKYLNFTSEKETSVEVHRFQVNNLQGR 973  
 QY 949 SLPLSLVLPVRLNQIVDRQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNVCS 1008  
 DB 974 DVPVSINFWPIELKGEAVW-TVMVSHQNPFLTQCYRNRUKPTQFDLLTHMQKSFVLDCS 1032

```

QY 1009 IAVCORIQOCDIPFGIOEBFNATLKGNSLDFWIKTSHNHLITVSTAEILENDVSFTLLP 1068
Db 1033 IADCLHLRCIDPGLGIDELYFLKGNLGFMTISQTLQKVKLLSBAEITFNTSVSYQLP 1092
QY 1069 GOGAFVRSQTETKVEFPNPLPLIVSGSVGLLILALITAAALYKLGFFKROKXDMGSE 1128
Db 1093 GQEAFLRAQTKTVLEMYKVHNPVPLIVSGSVGLLILALITAAALYKAGFFKROKXEMLEE 1152

RESULT 10
ITAD RAT STANDARD; PRT; 1161 AA.
AC Q9QVE7;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor.
GN Name=Itgad;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA O'Brien M.M., Vandervieren M., Kilgannon P.D., Dietsch G.,
RA Gallatin W.M.;
RT Cloning of rat alpha D, a novel beta 2 integrin.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and
CC VCAM1. May play a role in the atherosclerotic process such as
CC clearing lipoproteins from plaques and in phagocytosis of blood-
CC borne pathogens, particulate matter, and senescent erythrocytes
CC from the blood (By similarity).
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D
CC associates with beta-2 (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF021334; AAF21241.1; -.
DR HSSP; P11215; 1BHQ.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_alpha; 1.
DR Pfam; PF00592; VWFA; 1.
DR PRINTS; PR01185; INTEGRINA.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 5.
DR SMART; SM00327; VWFA; 1.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE; PS00234; VWFA; 1.
KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;
KW Repeat; Signal; Transmembrane.
FT SIGNAL 1 19 Potential.
FT CHAIN 20 1161 Integrin alpha-D.
FT DOMAIN 20 1100 Extracellular (Potential).
FT TRANSMEM 1101 1121 Potential.
FT DOMAIN 1122 1161 Cytoplasmic (Potential).
FT REPEAT 34 87 FG-GAP 1.

```

```

FT REPEAT 88 ?
FT DOMAIN 152 334 VWFA.
FT REPEAT 352 402 FG-GAP 3.
FT REPEAT 403 454 FG-GAP 4.
FT REPEAT 455 517 FG-GAP 5.
FT REPEAT 519 577 FG-GAP 6.
FT REPEAT 582 634 FG-GAP 7.
FT CA_BIND 467 475 Potential.
FT CA_BIND 531 539 Potential.
FT CA_BIND 594 602 Potential.
FT SITE 1126 1130 GPFKR motif.
FT DISULFID 69 76 By similarity.
FT DISULFID 108 126 By similarity.
FT DISULFID 656 711 By similarity.
FT DISULFID 769 775 By similarity.
FT DISULFID 845 860 By similarity.
FT DISULFID 993 1017 By similarity.
FT DISULFID 1022 1027 By similarity.
FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984A705E CRC64;

Query Match 55.2%; Score 3244.5; DB 1; Length 1161;
Best Local Similarity 57.6%; Pred. No. 1.2e-205;
Matches 649; Conservative 163; Mismatches 302; Indels 13; Gaps 9;

QY 2 NLDNTAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANQGSIIYQCDYGTGSCPIR 61
Db 21 NLDVEEIVFREDAASFGTVVFGSRLVGCAPLEAVAVNQTGRLYDCAPATGMCPIV 80
QY 62 LQVPEAVNMSLGLSLAATTSPPQLACGTPVHQTCSENYYVKGCLFPGSNLRQOPKF 121
Db 81 LRSPLAEAVNMSLGLSLVATNNAQLLACGTPAQACVKMYAKGSCLLIGSSL-QFIQAV 139
QY 122 PEALRGCEQSDSDIAPLVDSGSIIPHPERRAKEFIISTVMEQKKSKTSLFSLMOYSEER 181
Db 140 PASMECEPRQEMDIAPLVDSGSIINQDPAQKMFVKALMGEPASTSTLFLSMQYSNLIK 199
QY 182 IHFTKEFQNNPNRSLIKPITQLLGRTHATGIRKRVRELFNTNGARKNAFKILILIT 241
Db 200 THFTTEFKNILDPSLVDPVQLQGLIYATGIRTVMEELFHSKNGSRKSAKILLVIT 259
QY 242 DGEKFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQBELNTVASKPPRDHVPQINN 301
Db 260 DGQKYRDPLEYSDDVIPAADKAGIIRYAIVGVGDAFQEPALKEINTIGSAPPQDHVPKVG 319
QY 302 PEALATIONOLEKIFAIEGTGTGSSSFEHMSQSGSAATTSNGPLISTVGSYDWAG 361
Db 320 FAALRSIQRLQEKIFAIEGTQSSRSFQHEMSQEGFSALTSDGVLAGVGSFWSGG 379
QY 362 VFLYTSKSKSTFINMTRVDSMDNDAYGYAAAILLRNRVQSLVGLGAPRYOHLGLVAMPR 421
Db 380 AFLYPNTPEFFINNSQENVDMRDSYLGSTAVAFWKGVHSLILGAPRHOHTGKVIFITQ 439
QY 422 NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEYTRGQGVSCVPLP 481
Db 440 EARHWRPKSEVRGTQIGSYFGASLCSVDVDRDGSDDLVLIGAPHYYEYTRGQGVSVFVP 499
QY 482 RCQARWQCDVAVLYCEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYPH 541
Db 500 -GVRGRWQCEATLHCEQHPWGRFGVALTVLGDVNGDNLADVAIGAPGEESRGAIVYFH 558
QY 542 GTSGGIGSPSHSORIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQP 601
Db 559 GASRLINPSPQRTVSGQLSLRLQYFGQSLGGQDLTQDGLVDLAVGAQGHVLLRSPL 618

```

QY 602 VLKVKALMEENPREVARNVFECNDVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVTV 661  
Db 619 LLKVELSIRFAPMEVAKAVYQCHERTVLEAGEATVCLVHKVSPDLL--GNVQGSVRY 676  
QY 662 DLALDGRPHSRVFNFTKNSRRTQVGLTQTCETLKLQPLNCIBDPVSPVLRNLFNS 721  
Db 677 DLALDPCRLISRAIFDETCKNTLGRKTLGLGDHCETVKLLLPDCVEDAVSPIILRLNFS 736  
QY 722 LVGTPLSAGFNLRPVLAEADQRLFTALFPPEKNCNGNDNICODLSITFSPMSLDCLVVG 781  
Db 737 LVDRSASP--NLNHPVLAVGSDHTASLPPEKNCQKELLCEGDLGTSFNSGLQVLVVG 795  
QY 782 PREFNVTIVVRNDCGDSYRTOVTFPFLDLSYKRVSTLQORSQSWRLACESASSSTVS 841  
Db 796 SPELTIVTVVWNEGDSYGLTVKPYYPAGLSYRVTGTQ-QPHOYPLRLACEAPEAAQED 854  
QY 842 GALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANKVTSNNMPTNKTEFQ 901  
Db 855 --LRSSCSINHPIFREGAKTFTMTTFDVSYKAFGLGRLLLRKASSENKPDNTKTAQ 912  
QY 902 LEIPVKYAVVYVTSHGVSVKYLNFTASENTSR-VMOHQYQVSNLQORSPLISLVELVPV 960  
Db 913 LEIPVKYVTVLTSRQEDSTNHNFSHSGRRQEAHRYRVNLSPLKLA RVNFWVPV 972  
QY 961 RLQOTVLDWRPQVTFSENLS--TCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDI 1019  
Db 973 LINGVAWD---VTLSSPAQGVSCVSKMPQNPDPFLTQIRRSVLDCSIADCLHFRCDI 1029  
QY 1020 PFGIOEENFATLKNISFDWYIKTSHNHLIVSTABILFNDVSFTLLPQGGAFVRSQTE 1079  
Db 1030 PSLDIOELDFILRNLGSLFVSGVSTLQEKVLLVSEAITFTDTSVYSQLPQGEAFLRAQVE 1089  
QY 1080 TKVEPEVPNPLPLIVGSSVGGLLILALITAAVYKLGPFKRYKQVDM 1126  
Db 1090 TLEEVYVYBPIFLVAGSSVGGLLILALITVLYKLGFFKRYKQVEM 1136

## RESULT 11

ID - ITAL BOVIN STANDARD; PRT; 1165 AA.  
AC P61625;  
DT 05-JUL-2004 (Rel. 44, Created)  
DT 05-JUL-2004 (Rel. 44, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1  
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha  
DE chain) (CD11a).  
GN Names:ITGAL;  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA PubMed=14697514; DOI=10.1016/j.gene.2003.09.043;  
RA Fett T., Zechinon L., Baise E., Deamecht D.;  
RT "The bovine (Bos taurus) CD11a-encoding cDNA: molecular cloning,  
RT characterisation and comparison with the human and murine  
RT glycoproteins";  
RL Gene 325:97-101(2004).  
CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,  
CC ICAM3 and ICAM4. It is involved in a variety of immune phenomena  
CC including leukocyte-endothelial cell interaction, cytotoxic T-cell  
CC mediated killing, and antibody dependent killing by granulocytes  
CC and monocytes (By similarity).  
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L  
CC associates with beta-2 (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
CC with I-domains do not undergo protease cleavage.  
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.

CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
CC -!- SIMILARITY: Contains 1 VWFA domain.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).

CC -----  
CC EMBL; AY267467; AAP94035.1; -.  
CC PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
CC DR PROSITE; PSS0234; VWFA; 1.  
CC KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
KW Repeat; Signal; Transmembrane.  
FT SIGNAL 1 23 Potential.  
FT CHAIN 24 1165 Integrin alpha-L.  
FT DOMAIN 24 1084 Extracellular (Potential).  
FT TRANSMEM 1085 1105 Potential.  
FT DOMAIN 1106 1165 Cytoplasmic (Potential).  
FT REPEAT 40 88 FG-GAP 1.  
FT REPEAT 89 146 FG-GAP 2.  
FT REPEAT 153 324 VWFA.  
FT REPEAT 347 397 FG-GAP 3.  
FT REPEAT 398 452 FG-GAP 4.  
FT REPEAT 454 513 FG-GAP 5.  
FT REPEAT 515 572 FG-GAP 6.  
FT REPEAT 575 627 FG-GAP 7.  
FT CA BIND 465 473 Potential.  
FT CA BIND 527 535 Potential.  
FT CA BIND 587 595 Potential.  
FT SITE 1111 1115 GFFKR motif.  
FT DISULFID 71 78 By similarity.  
FT DISULFID 108 126 By similarity.  
FT DISULFID 650 704 By similarity.  
FT DISULFID 768 774 By similarity.  
FT DISULFID 842 858 By similarity.  
FT DISULFID 994 1009 By similarity.  
FT DISULFID 1017 1048 By similarity.  
FT CARBOHYD 33 33 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 86 86 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 185 185 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 646 646 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 667 667 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 723 723 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 859 859 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 894 894 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 929 929 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 1056 1056 N-linked (GLCNAC..) (Potential).  
FT CARBOHYD 1067 1067 N-linked (GLCNAC..) (Potential).  
SQ SEQUENCE 1165 AA; 128725 MW; DAEB3A3F1E1463CB CRC64;

Query Match 26.9%; Score 1583; DB 1; Length 1165;  
Best Local Similarity 35.1%; Pred. No. 1.4e-95;  
Matches 415; Conservative 195; Mismatches 466; Indels 106; Gaps 33;

QY 1 FNLDTENAMTFQENARG--FGQSVVLQGSRVVVGAPQEIIVAAQNRSLYQCDYSGSC 58  
Db 24 YNLDFVRHVQNFSPFLAGRHFGRVVLQV-GNVVVVGAPSE---GNSMNLVQCQFETGDC 79  
QY 59 PIRLQVPEAVNMSLGLSLAATSPOLLACGPTVHTQTSENTYVVKGLCFPLGNSNLRQQP 118  
Db 80 PVLTS--SNVTSKYLGMTLATDFTSDNLLACDPGLSRTCDQNIYLSGLCVLIHENLRGPV 137  
QY 119 OKFPEALRGCPEDSDIAFLVDGSGSIIPHDPRRAKEFTISVMEQLKKSKTILSLMOYSE 178  
Db 138 LQHPGQYQECIKGNVDLVFLFDGMSLQQDEFEKIVDFMKDKVKKLSNSSYQFAAVQFST 197  
QY 179 EFRHTEFKEFQNNPNRSLIKETITQLGRTHATGIRKVVRELFNITNGARKQAFKILI 238  
Db 198 YFRTEFTFLDIIRKOPDALLAGVKHRLTLNTFGAINVYAKEVFRFDLGRPDATKVL 257



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; Y00796; CAA68747.1; --  
DR EMBL; AC002310; AAC31672.1; --  
DR PIR; S03308; S03308.  
DR PDB; 1CQP; X-ray; A/B=153-334.  
DR PDB; 1DGO; NMR; A=149-336.  
DR PDB; 1LFA; X-ray; A/B=--  
DR PDB; 1MJN; X-ray; A=153-331.  
DR PDB; 1MQ8; X-ray; B/D=155-331.  
DR PDB; 1MQ9; X-ray; A=152-330.  
DR PDB; 1MQA; X-ray; A=152-330.  
DR PDB; 1ZON; X-ray; @=150-336.  
DR PDB; 1ZOO; X-ray; A/B=150-336.  
DR PDB; 1ZOP; X-ray; A/B=150-336.  
DR Genew; HGNC:6148; ITGAL.  
DR MIM; 153370; --  
DR GO; GO:0008305; C:integrin complex; TAS.  
DR GO; GO:0006928; P:cell motility; TAS.  
DR InterPro; IPR00413; Integrin\_alpha.  
DR InterPro; IPR002035; VWF\_A.  
DR Pfam; PF01839; FG-GAP; 3.  
DR Pfam; PF00357; Integrin\_alpha; 1.  
DR Pfam; PF00092; VWA; 1.  
DR PRINTS; PR01185; INTEGRINA.  
DR PRINTS; PR00453; VWFADOMAIN.  
DR SMART; SM00191; Int\_alpha; 5.  
DR SMART; SM00327; VWA\_1.  
DR PROSITE; PS00242; INTEGRIN ALPHA; 1.  
DR PROSITE; PS50234; VWFA; 1.  
KW 3D-structure; Alternative splicing; Calcium; Cell adhesion;  
KW Direct protein sequencing; Glycoprotein; Integrin; Magnesium;  
KW Receptor; Repeat; Signal; Transmembrane.  
FT SIGNAL 1 25  
FT CHAIN 26 1170 Integrin alpha-L.  
FT DOMAIN 26 1090 Extracellular (Potential).  
FT TRANSMEM 1091 1111 Potential.  
FT DOMAIN 1112 1170 Cytoplasmic (Potential).  
FT REPEAT 42 91 FG-GAP 1.  
FT REPEAT 92 149 FG-GAP 2.  
FT DOMAIN 156 327 VWFA.  
FT REPEAT 350 400 FG-GAP 3.  
FT REPEAT 401 455 FG-GAP 4.  
FT REPEAT 457 516 FG-GAP 5.  
FT REPEAT 518 575 FG-GAP 6.  
FT REPEAT 578 630 FG-GAP 7.  
FT CA\_BIND 468 476 Potential.  
FT CA\_BIND 530 538 Potential.  
FT CA\_BIND 590 598 Potential.  
FT SITE 1115 1119 GFFKR motif.  
FT DISULFID 73 80 By similarity.  
FT DISULFID 111 129 By similarity.  
FT DISULFID 653 707 By similarity.  
FT DISULFID 771 777 By similarity.  
FT DISULFID 845 861 By similarity.  
FT DISULFID 998 1013 By similarity.  
FT DISULFID 1021 1052 By similarity.  
FT CARBOHYD 65 65 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 89 89 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 188 188 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 649 649 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 670 670 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 726 726 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 730 730 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 862 862 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 885 885 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 897 897 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 1060 1060 N-linked (GlcNAc... ) (Potential).  
FT CARBOHYD 1071 1071 N-linked (GlcNAc... ) (Potential).

FT VARSPLIC 954 954  
FT FT  
FT FT  
FT CONFLICT 214 214  
FT CONFLICT 660 660  
FT STRAND 155 162  
FT STRAND 164 164  
FT TURN 165 166  
FT HELIX 169 185  
FT TURN 186 188  
FT STRAND 191 198  
FT STRAND 202 206  
FT STRAND 208 214  
FT HELIX 217 221  
FT TURN 222 223  
FT STRAND 229 229  
FT HELIX 233 243  
FT TURN 244 244  
FT HELIX 247 249  
FT TURN 250 250  
FT TURN 253 254  
FT STRAND 256 263  
FT HELIX 274 276  
FT TURN 277 278  
FT STRAND 280 286  
FT HELIX 288 290  
FT HELIX 293 297  
FT TURN 298 299  
FT HELIX 300 302  
FT HELIX 307 310  
FT STRAND 311 314  
FT HELIX 317 319  
FT HELIX 323 330  
SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;  
Query Match 26.5%; Score 1557.5; DB 1; Length 1170;  
Best Local Similarity 34.7%; Pred. No. 7.1e-94;  
Matches 407; Conservative 206; Mismatches 460; Indels 101; Gaps 37;  
Qy 1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGLVQCDYSTGSC 58  
Db 26 YNLDRVGRSFPSPRAGRHFGRVLIQV-GNGVIVGAPGE---GNSTGSLVQCQSGTGCHL 81  
Qy 59 PIRLQVPVEAVNMSLGLSLAATTPPOLLAGCPVTHOTCSNTYVVKGLCFLFGSNLR--- 115  
Db 82 PVLIR-GSNVTSKYLGMTLATDPTDGSILACDPGLSRTCDQNTYLSGLCYLFRQNLQGM 140  
Qy 116 -QQPKFPEALRGCPQSDSIAFLVDGSGSIIPHDFRAKEFIISTVMEQLKKSKTLP 174  
Db 141 LQGRPGFOECIKG---NVDLVFLFDGMSLQPDPEFQKILDFMKDVMKLSNTSYQFAAV 196  
Qy 175 QYSEFRIHFTFEFQNNPNRSLIKPITOLLGRTHTATGIRKVVRELFNITNGARKNAF 234  
Db 197 QFSTSYKTEFDSDYVVRKOPDALLKHVGHMLLTNTFGAINVYATEVFRRELGARPDAT 256  
Qy 235 KILILITDGEKFGPLGVEDVPEADREGVIRYVIGVGDAFRSEKSRQELMTVASKPPRD 294  
Db 257 KVLIIITDGE--ATDSGNIDAADK-----IIRYIIIGIKHQTKESETLHKFKSPASE 309  
Qy 295 HVFOINNFEALKTIONLREKIPIAETGTQTSSSPHEMSQEGFSAAITNSGPLLSTVG 354  
Db 310 FVKILDTFEKLKDLTFELQKKIYVIEGTSKDLSFNMELSSSGISADLSRGHVAVVGAVG 369  
Qy 355 SYDNAGGVF-LYTSKEKSTFNMTREVDSDMNDAYLVGAAA-IILNRVQSLVLGAPRYOH 412  
Db 370 AKDWAGGFLDLKADLQDDTFIENBPLTEVRAVGLYGVTVWLPDRQKTSLLASGAPRQ 429  
Qy 413 IGLVAMPR--QNTGMWESNANVKGTQICAYGASLCSVDVDSNGSTDLVLICAPHYEQ 470  
Db 430 MGRVLLFQEPQGGHWSQVQTIHGTQIGSYFGGELCGVDVDQDGETELLIGLAPLFYGEQ 489  
Qy 471 RGGQSVCPPLRQGRARWQCDAV--LYGECQGPWRFGAALTVLGDVNGDKLTDVAIGAP 528

Db 490 RGRVFYIY-----QRRQLGPEEVSELDQPGYPLGRGPEAITALTIDINGDLVDVAVGAP 544  
 QY 529 GEEDNRGAVLYEGTSGSISPSHSORLAGSKLSPRLQYEGQSLSGGQDLTMDGLVDLTV 588  
 Db 545 LEE--QRAVIFNGRHG-GLSPQPSQRIGETQVLSIQWFGRSHGVKDLGGGLADVAV 601  
 QY 589 GAQGHVLLRSQPYLRVKAIMEFNPREVARNVFECNDQVY-KGKEAGEVRVCLHVKQKSTR 647  
 Db 602 GAESQMVLSRPPVDMVMTLMSFSPAIPVHEVECSYSTSNKMKEGVNIICQI-KSLY 660  
 QY 648 DRLREGIQSVVYDIALDGRPHSRVAFNETKNSTRQTOVLGLTQTCTETLKLQLPNCI 707  
 Db 661 PQF-QCGLVANLTYTLQDGHRTRRRGLFPGGRHELRLNAVT-TSMSCDFFSHPDVCV 718  
 QY 708 EDVPSVILRLNLSL---VGTPLS--AFGN-----LRPVLAEDAQRLLFTALFPFEKNCGN 757  
 Db 719 QDLISPINVLSLMSLEEETPRDQAGKDIPILRPSLSHSTWET-----PFERKNCGE 773  
 QY 758 DNICQDLSITFSFMSLDCLVGGPPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVS 817  
 Db 774 DKCEANLRVSPARSRLRLTAFASLSVELSLNLEEDAYVQLDLHPPLGSPKVE 833  
 QY 818 TLQNRORSWRILACES--ASSTEVGALKSTGCSINHPIPPENSEVTFNITFDVDSKAS 875  
 Db 834 ML---KPHSQIPVSCPELPEESRLSRAAL---SCNVSSPIFKAGHSVALQMMFNTLVNS 887  
 QY 876 LGNKLKLLKANVTSENN---MPTNKTETFOLELPVKYAVVMVTVSHGVSTKYLNFNTASEN 931  
 Db 888 WGSVVELHANVTNNEDSLLSENSATTI---IPILYIPINILQDQEDSTLYYSFYPKGP 944  
 QY 932 TSKVMQHOYQV---SNLQORSLP-LSLVLVVPVRLNQTQVTFNTPQVTFSENLSTCHTK- 986  
 Db 945 KIHQVKHMQVRIQPSIHHDNIPTLEAVGVQPPSEGPITHQWSVQMEPPV--PCHYED 1002  
 QY 987 -BRLPHSD--FLAELRKAPVNCSTAVCQRIQCDIPFGIQIEFNATLKNLSFDWYIK 1043  
 Db 1003 LERLPDAEPPCLPGALFRCPVW-----FRQELLVQVIGTLELVGEIE 1044  
 QY 1044 TSHNHLIIVSTABILENDSVFTLLPGQAFVRVSTQETKVEPFPVNPPLPLIVGSSVGLL 1103  
 Db 1045 AS-SMFSLCSSLSISNSKSHFLYGSNASL-AQVVMKVDVVYEQWMLYLVLSGGIGLL 1102  
 QY 1104 LLAALTAALYKLGFFRKQYKDMMSSEG-GPPGAP 1136  
 Db 1103 LLLLIIFVLYKGVFFKRLNKKMEAGRVNGIP 1136  
 RESULT 13  
 Q6TYB8  
 ID Q6TYB8 PRELIMINARY; PRT; 1166 AA.  
 AC Q6TYB8;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 SEQUENCE FROM N.A.  
 RA Dileepan T., Thambikat P., Kannan M.S., Maheswaran S.K.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 CC -! SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -! SIMILARITY: Belongs to the integrin alpha chain family.  
 DR EMBL; AY382558; AAQ90015.1; -  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR02035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 1.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWA; 1.  
 DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 KW Cell adhesion; Integrin; Transmembrane.  
 SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;  
 Query Match 26.5%; Score 1557; DB 2; Length 1166;  
 Best Local Similarity 34.6%; Pred. No. 7.6e-94;  
 Matches 409; Conservative 198; Mismatches 469; Indels 106; Gaps 33;  
 QY 1 ENLDTENAMTFOENARG--PGQSVVQLQGRSVVVGAPQEIIVAAANQRGSLYQCDYSTGSC 58  
 Db 25 YNDLVHVHVNQFSPPLAGRHFGYRVLOV-GNVVVVGAPSE---GNSMGNLYQCQFETGDCL 80  
 QY 59 PIRLOVPEAVNMSLGLSLAATTPFQLLACGPTVHTQCTSENTYVKGKLCFLFGSNLRQOP 118  
 Db 81 PVTLS--SNVTSKYLQMTLATDFTSNLACDCLSRTCQNIYLSGLCVLIHENLRGPV 138  
 QY 119 QKPEPALRGCPQSDSIAFLVDGSGSIIIPHDFFRAKEFISTVMEQLKKTFLFSLMOYSE 178  
 Db 139 LQHPGYQECIKGNVDLVELFDGMSLQODEFEKIVDFMKDVMKLSNSGYQFAAQFST 198  
 QY 179 ERIHETEKEFQNNPNSRLIKPITQLLGRTHATGIRKVVRELFNITGARKNAFKILI 238  
 Db 199 YRTEETFDLYIKQKDPDALLAGVHKMRLTNTFGAINYVAKVEVFRPDILGARDATKVI 258  
 QY 239 LITDGKFGDPLGYEDVPEADREGVIRYVGVGDAPRSEKRSQELNTVASKPRDHVFQ 298  
 Db 259 IITDGK--PPTNTTLMRPKTSS---RSLLGIGKNFKTESQEAHQFASKVEFEVKI 311  
 QY 299 INNFEALKTIQNLREKIPAIETQGTGSSSPHEMSQBFSAITSNGLSTVGSYDW 358  
 Db 312 LDTFEKLKDLFTLQKIVIEGTSKODLTSFMNELSSGISADLSEGHGWGAVGAKDW 371  
 QY 359 AGSVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAA-IILNRVQSLVIGAPRYOHIGLV 416  
 Db 372 AGFLDLKADLSSTVGNELTVESRAGLYGVTVTRLPSSRGTMSSLATCAPKYOHVGRV 431  
 QY 417 AMFRQ--NTGMWSEANANVGTOIGAYFGASLCSVDVDSNGSTDVLVLIGAPHYYEOTRGQ 474  
 Db 432 LLFQPKRGKQPSQIEIDGIGSYFGGELCGVDVDRDGETELLIAIAPLYGEGRGR 491  
 QY 475 VSVCLPRQARMQCDVLYGEGQGWGRFGAALTGLGVNGDKLTDVAIGAPEGEDNR 534  
 Db 492 VFIY---QXIQLEFQWVSELOGETGYPLGRFGAAIAALTIDINGDELTDVAVGAPLEE--Q 546  
 QY 535 GAVYLFHGTSGSGISPSHSORIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHV 594  
 Db 547 GAVYIFNGQQG--GLSPRPSQRIGETQVLSIQWFGRSHGVKDLGGGLADVAVGAGQV 605  
 QY 595 LLLRSQPYLRVKAIMEFNPREVARNVFECNDQVWKGKEAG-EVRVCLHVQK--STRDLR 651  
 Db 606 IVLSRPPVDIITSVSPSPAIPVHEVECSYSTSNKMKEGVNLTVCFQVKSLSLT----F 661  
 QY 652 EQIQSVVYDIALDGRPHSRVAFNETKNSTRQTOVLGLTQTCTETLKLQLPNCIEDPV 711  
 Db 662 QGHLVANLTYTLQDGHRTRRRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHPFICIDLI 720  
 QY 712 SPVILRLNLSL---VGTPLS--AFGNLRPVLAEDAQRLLFTALFPFEKNCGNINICQDLS 766  
 Db 721 SPINVLSYSLMEEETPRDPRALDRDIPILKPSHLETKEIPFEKNCGEKNCDEADLK 780  
 QY 767 ITFSFMSLDCLVGGPPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSR 826  
 Db 781 LAFSDMRSKILLTPSASISVLETLRNTAEADAYVQVTLSPFQGLSFRKVEL---KPHS 837  
 QY 827 SWRLACESASSTEVGALKSTGCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANV 886  
 Db 838 HVPVGCCELPEEAVVHS-RALSCNVSSPIFGEDSMVDIQVMFNTLQKSGWGFIELQANV 896  
 QY 887 TS-----ENNMPRTNKTBFQLELPVKYAVVMVTVSHGVSTKYLNFNTASENTRVMOH 938

```

Db      897 SCNNEDSLLLEDNSATTS-----IPVMTPIINVLTQKQENSTLYISFTPKSPRIHVHVK 949
QY      939 QYQV-----SNLQORSPLISLVFLVPLVRLNQTVI---WD---RPQVTFSS-ENLSSTCHTKE 987
Db      950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPNLESDEAE 1008
QY      988 RLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIPFGIOEFNATLKNLSFDWIKTSHN 1047
Db      1009 -----SCSFGT--EFRCPIDF---RQBILVQVNGMVELRGTIKAS-S 1044
QY      1048 HLLIVSTAELFENDSVFTLLPGOGAFVRSOTETKVEPEFVNPPLIVGSSVGLLALL 1107
Db      1045 MLSCSSLAISFNSSKHFHLHGRNASM-AQVWVKVDIVYKEMLYLYVLSIGGLLLFL 1103
QY      1108 ITAALYKLGFFKQYKDDMM-----SEGGPFGBFPQ 1137
Db      1104 IFIALLYKVGFFKRLNKEKMEANVDASSEIPGEDAGQPELEKE 1145

RESULT 14
AAQ90015 PRELIMINARY; PRT; 1166 AA.
AC AAQ90015;
DT 02-MAR-2004 (TrEMBLrel. 27, Created)
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;
RT "Molecular cloning and sequencing of bovine CD11a.";
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY382538; AAQ90015.1; -
SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;

Query Match 26.5%; Score 1557; DB 2; Length 1166;
Best Local Similarity 34.6%; Pred. No. 7.6e-94;
Matches 409; Conservative 198; Mismatches 459; Indels 106; Gaps 33;

QY      1 FNLDTENAMTFQENARG--FQGSVVQLQGSRRVTVGAFQEIIVAAQNRSLYQCDYSTGSCSE 58
Db      25 YNLDRVHVQNFSPFLAGRHFGRYVLQV--GNGVVVVGAPSE---GNSMGNLVQCQPETGDC 80
QY      59 PIRLOVPEAVNMSLGLSLAATTSPPOLLAGCPVTHQTCSENTYVKGICFLFGSNLRQOP 118
Db      81 PVTLSS--SNYTSKYLGMVTLATDPTSDNLLACDPLSKTCDQNIYLSGLCYLIHNLGRPV 138
QY      119 QKPEALRGCPQEDSIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKKSKTLFSLMOYSE 178
Db      139 LQHPGQVCEIKGNVDLVFLFDGSMISLQODEFEKIVDFMKDVMKLSNYSYQFAVQFST 198
QY      179 EFRTHFTFKFPQNNPRLSIKPTQLLGRTHATGKRVKRVRELFNTNGARKNAFKILI 238
Db      199 YFRTEFTFLDYIKQDPDALLAGVKHMLLNTTFGAINYVAKEVFRPDLGARPDATKVI 258
QY      239 LITDEKFGDPLGVEDVPEADREGVIRYVGVGDAPFSEKSRQELNTVASKPPRDHVQ 298
Db      259 IITDCK---PPTNTILMRPKTSS-----RSLGIGKNFKTESQALHQPASKPVEEFVKI 311
QY      299 INNFEALKTIONLREKIFAIEGTQGTGSSSFEHMSQEGFSAATISNGPLSTVGSYDW 358
Db      312 LDTPEKLDLTFELQKKIYVIEGTSKQDLTFSFNMELSSGSIADLSSEHGTVGVGAVGAKDW 371
QY      359 AGGVF-LYTSKEKSTFINNTRVDSMDNDAYLGAAA-IILNRRVQSLVILGAPRVQHIGLV 416
Db      372 AGGFLLDKADLKSTFTVGNELPTVESRAGYLGIVTVTRPLPSRGTMSLLATGAPKYQHVGRV 431

```

```

QY      417 AMFRQ--NTGMWBSNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQ 474
Db      432 LLFQQPKRGPMQSQIIEIDGIGISVYFGGFLCGVDVDRDGETELLIIAALYYIGEQRGR 491
QY      475 VSYCPILPRGORARWQCDAVLYGEGQPMWGFAGALTVLVGVNGDKLTDVAIGAPGEDNR 534
Db      492 VFIV---QKIQLEBFQVMVSELOGETGYPLGRFGAIAALTDINGDELTDVAVGAPLEF--Q 546
QY      535 GAYLPHGTSGSIGSPHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAOGHV 594
Db      547 GAYIIFNGQGG--GLSRPFSQRIEGTFWFGSIHGVKDLGGGLADVAVGAGQV 605
QY      595 LLRSOPVLRVKAIMFENPREVARNVPECNDQVVKGKEAG-EVRVCLHVOK--STRDLRL 651
Db      606 IVLSRPVVDIITVSFSFAEIPVHEVECSYSTSNQKKEGVNLTVCFQVKSLLST----F 661
QY      652 EGOIQSVVTVYDLALDSGRPHSRVAFNETKNSTRQTVGLGLTOTCETLKLQLENCHDPV 711
Db      662 QGHLVAMLVTVTLQDGHRTSRGLFPFGKHLIGNTAVTPV-KSCFVFWFHFHFCIQDLI 720
QY      712 SPIVLRNFSL---VGTPLS--AFGNLRPVLAEDAQRLFTALPPEFKNCGNDNICQDDLS 766
Db      721 SPINVSLSYSLWEEEGTPDRALDRDIPILPKSPHLETKYKIFFEKNCGEDKNCEADLK 780
QY      767 ITFSFMSLDCLVVGGPREFNVTVTVNRDGEDSYRTQVTFEFLDLSYRKVSTLQNRQSR 826
Db      781 LAFSDMRSKTLRLTPSASLSVRLTLRNTAEDAYVQVTLSPQGLSPFRKVEIL--KPHS 837
QY      827 SWRLACESASTEVSGALKSTCSINHIPPENSEVTFNITFDVDSKASLGNKLLKLVANV 886
Db      838 HVPVGCBELPEEAVVHS-RALSCNVSSPIFGESMVDIQVMFNTLQKSGWGDFTLEQANV 896
QY      887 TS-----ENNMPRTNKTEFQLELPVYAVVMTVSHGVSTKYLNFTASENTRVWQH 938
Db      897 SCNNEDSLLLEDNSATTS-----IPVMTPIINVLTQKQENSTLYISFTPKSPRIHVHVK 949
QY      939 QYQV-----SNLQORSPLISLVFLVPLVRLNQTVI---WD---RPQVTFSS-ENLSSTCHTKE 987
Db      950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPNLESDEAE 1008
QY      988 RLPSHSDFLAELRKAPVWNCISIAVCQRIQCDIPFGIOEFNATLKNLSFDWIKTSHN 1047
Db      1009 -----SCSFGT--EFRCPIDF---RQBILVQVNGMVELRGTIKAS-S 1044
QY      1048 HLLIVSTAELFENDSVFTLLPGOGAFVRSOTETKVEPEFVNPPLIVGSSVGLLALL 1107
Db      1045 MLSCSSLAISFNSSKHFHLHGRNASM-AQVWVKVDIVYKEMLYLYVLSIGGLLLFL 1103
QY      1108 ITAALYKLGFFKQYKDDMM-----SEGGPFGBFPQ 1137
Db      1104 IFIALLYKVGFFKRLNKEKMEANVDASSEIPGEDAGQPELEKE 1145

RESULT 15
ID ITAL MOUSE STANDARD; PRT; 1163 AA.
AC P24063;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Name=Itgal; Synonyms=Lfa-1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91268576; PubMed=2051027;
RA Kaufmann Y., Tseng E., Springer T.A.;
RT "Cloning of the murine lymphocyte function-associated molecule-1

```





```

Db 716 IQDLISFVNSLSEESTPRDQGRAMQPIRPSIHTV-TKEIPFEKNGEDKKCE 774
Qy 763 DDLISITFSFMSLCLVVGGP-----REFNVTVTRNDGEDSYRTQVTFPPLDLSYRKV 816
Db 775 ANLTLSPPARS-----GPLRLMSSASLAVETWLSNGEDAYWVRDLDFPRGLSPRKV 827
Qy 817 STLQNRQSRWRLACHESASTEVSGAL-KSTSCSINHPIPPENSEVTENITFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCBEL--TEGSSLLTKLKNVSSPIFKAGQEVSLQVMFNTLLNS 882
Qy 876 LGNKLILKANVTSEN-NMPTNKTEFOLELPVKYAVYVTVTSHGVSTKYLNFTASENISR 934
Db 883 WEDFVELNGTVHCENENSQEDNSAATHIPVLPVNLTKQENSTLYISFTPKGPKTQ 942
Qy 935 VMQHQQVSNLQGRSLPISLVLVPLVRLNQTVIWRPQ----VTFSENLS-----TCHTK 986
Db 943 QVCHVYQV-----RIQPSAYDHNMT-LEALVGVPRPHSEDLITYTWSVQTDPLVICHSE 996
Qy 987 E-RLPSHSDFLAELRKAPVYVNCSTAVCQRIQCDIPFFGQIEEFNATLKGNSLSPDWYIKTS 1045
Db 997 DLKRPSE---AEQPCLPV-----QFRCPIVF---RWEILIQTVGVLSKEIKAS 1042
Qy 1046 HNHLLIVSTABILFNDVSFTLLPGQGAFFVRSQTETKVEPEVPNPPLPLIVGSSVGGLLLL 1105
Db 1043 -STLSLCSLSVSFNSKHFHYGSKA-SEAQLVAVKVDLIHEKEMLHVYVLSGIGGLVLL 1100
Qy 1106 ALITAAALYKLGFFKRYQYKDM-SEGPPGAP 1136
Db 1101 FLIFLALYKVGFFKRNLRKMEADGGVNGSP 1132

```

Search completed: November 9, 2004, 12:17:38  
Job time : 141.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds  
(without alignments)  
3950.365 Million cell updates/sec

Title: US-09-902-481b-5

Perfect score: 5876

Sequence: 1 ENLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPGGPAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5862	99.8	1153	2	AAW65090 Human Bet
2	5862	99.8	1153	3	AAW65090 Human Bet
3	5862	99.8	1153	5	AAU80252 Human int
4	5862	99.8	1153	5	ABG61469 Human Bet
5	5862	99.8	1153	5	AAO14428 Integrin
6	5862	99.8	1153	7	ADD25615 Binding d
7	5852	99.6	1153	2	AAW04136 Alpha sub
8	5846.5	99.5	1152	8	ADM99589 Human int
9	5839.5	99.4	1152	8	ADP12435 Protein e
10	3483	59.3	1163	8	ADP44061 Human CDI
11	3469	59.0	1163	8	ADN02004 Human inf
12	3469	59.0	1163	8	ADQ17510 Human sof
13	3460	58.9	1163	2	AAW07120 p150.95 a
14	3446	58.6	1163	2	AAW65091 Human Bet
15	3446	58.6	1163	3	AAW07361 Human CDI
16	3446	58.6	1163	5	ABG61470 Human Bet
17	3444	58.6	1163	6	ABU07406 Protein d
18	3444	58.6	1163	7	ADG32005 Human hom
19	3411	58.0	1161	2	AAW78166 Human bet
20	3411	58.0	1161	2	AAW23049 Human bet
21	3411	58.0	1161	2	AAW57491 Human bet
22	3411	58.0	1161	2	AAW65089 Human Bet
23	3411	58.0	1161	2	AAW72825 Human alp
24	3411	58.0	1161	2	AAW73342 Human alp
25	3411	58.0	1161	3	AAW07359 Human alp

## ALIGNMENTS

## RESULT 1

AAW65090

ID AAW65090 standard; protein; 1153 AA.

XX AC AAW65090;

DT 28-SEP-1998 (first entry)

XX Human Beta-integrin CD11b subunit protein.

DE Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;  
KW rheumatoid arthritis.

OS Homo sapiens.

PN US5728533-A.

XX 17-MAR-1998.

PD 07-JUN-1995; 95US-00485618.

XX 23-DEC-1993; 93US-00173497.

PR 05-AUG-1994; 94US-00286889.

PR 21-DEC-1994; 94US-00362652.

XX (ICOS-) ICOS CORP.

PI Van Der Vieren M, Gallatin WM;

XX WPI; 1998-206565/18.

XX Screening assay for modulators of integrin binding - using immobilised or

XX labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.

XX Example 5; Fig 1A-D; 106pp; English.

CC This sequence represents a human beta-integrin CD11b subunit which is  
CC used to describe a method for identifying compounds that modulate the  
CC interaction of the beta-integrin alpha-d subunit with a binding partner  
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
CC -d binding partner, one of which is immobilised and the other of which is  
CC labelled, in the presence of a test compound, and determining if the  
CC compound affects binding between the alpha-d polypeptide and alpha-d  
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
CC comprising the cytoplasmic, transmembrane or extracellular domain of  
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

ABG61468 Human Bet  
AAW23064 Human Bet  
AAW65106 Human Bet  
AAW72837 Human alp  
AAW73343 Human alp  
AAW07376 Human alp  
ABG61485 Human Bet  
ABG61485 Human Bet  
AAW78169 Rat alpha  
AAW23062 Rat beta  
AAW60004 Rat alpha  
AAW72824 Rat alpha  
AAW07374 Rat alpha  
ABG61483 Rat Beta2  
AAW65104 Rat beta-  
AAW73345 Rat alpha  
AAW23061 Mouse bet  
AAW60003 Mouse bet  
AAW65103 Mouse alp  
AAW72836 Mouse alp  
AAW73347 Mouse alp

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
CC and rheumatoid arthritis  
XX

SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 2; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDENAMTQENARGFQSVVQLQGSRVVVGAPQEIIVAAANRGSLYQCDYSTGSCPEI 60  
DB 17 FNLDENAMTQENARGFQSVVQLQGSRVVVGAPQEIIVAAANRGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLAGCTVTHQTCSENYYVKGCLFGLFSNLRQOPQK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTPPOLLAGCTVTHQTCSENYYVKGCLFGLFSNLRQOPQK 136  
QY 121 FPEALRGCPQSDSIAFLVDGSGSIIPHDFRAKEPISTVMEQLKSKTFLSLMOYSEEF 180  
DB 137 FPEALRGCPQSDSIAFLVDGSGSIIPHDFRAKEPISTVMEQLKSKTFLSLMOYSEEF 196  
QY 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELENTNGARKNAKILILI 240  
DB 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELENTNGARKNAKILIVI 256  
QY 241 TDGEKFGDPLGVIEDVPEADREGVIRYVIGVDAPFRSEKSEQLNTVASKPRDHVFOIN 300  
DB 257 TDGEKFGDPLGVIEDVPEADREGVIRYVIGVDAPFRSEKSEQLNTVASKPRDHVFOIN 316  
QY 301 NFEALKTTIONLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360  
DB 317 NFEALKTTIONLREKI FAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTVRVDSMDNAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTVRVDSMDNAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVCP 480  
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQSVCP 496  
QY 481 PRGQARWQCDVAVLXGEOQPMWREGAALTVLGDVNGDKLTDVAIGAPGEEDNRCGAVL 540  
DB 497 PRGQARWQCDVAVLXGEOQPMWREGAALTVLGDVNGDKLTDVAIGAPGEEDNRCGAVL 556  
QY 541 HGTSGSGISPHSQRSIAGSKLSPRLQYFQSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
DB 557 HGTSGSGISPHSQRSIAGSKLSPRLQYFQSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 616  
QY 601 PVLRYKATMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLREGQIQSVYT 660  
DB 617 PVLRYKATMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLREGQIQSVYT 676  
QY 661 YDLALDSGRPSRAVFNETKNSRTOVLGTOTCETLKLQLPNCIEDPVSPIVRLNLF 720  
DB 677 YDLALDSGRPSRAVFNETKNSRTOVLGTOTCETLKLQLPNCIEDPVSPIVRLNLF 736  
QY 721 SLVGTPLSAFNLRLVLAEDAQRLLTALPFPEKNCNDNIQDDLSITFSFMSDCLVVG 780  
DB 737 SLVGTPLSAFNLRLVLAEDAQRLLTALPFPEKNCNDNIQDDLSITFSFMSDCLVVG 796  
QY 781 GPRFNVTVRNDEGDSYRQVTFEFDLISYKRVSTLQORSORSELACESASTEV 840  
DB 797 GPRFNVTVRNDEGDSYRQVTFEFDLISYKRVSTLQORSORSELACESASTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRINKTEF 900  
DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRINKTEF 916  
QY 901 QLELPVKAVYVMVVTSHGVSTKYLNFNTASENTRVMQHQYQVSNLQORSLFISLVFLVPV 960  
DB 917 QLELPVKAVYVMVVTSHGVSTKYLNFNTASENTRVMQHQYQVSNLQORSLFISLVFLVPV 976

QY 961 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVWVNCIAVCQRIQCDDIP 1020  
DB 977 RLNOTVIWDRPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVWVNCIAVCQRIQCDDIP 1036  
QY 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAELIFNDSVFTLLPGQCAFVRSQDET 1080  
DB 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNHLLIVSTAELIFNDSVFTLLPGQCAFVRSQDET 1096  
QY 1081 KVEFEFVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKMMSEGGPPGAPQ 1137  
DB 1097 KVEFEFVNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKMMSEGGPPGAPQ 1153

RESULT 2

AAB07360

ID AAB07360 standard; protein; 1153 AA.

XX AC AAB07360;

XX AC AAB07360;

DT 17-JAN-2001 (first entry)

XX Human CD11b protein sequence.

DE Human; macrophage infiltration inhibition; alpha d integrin;

XX Human; macrophage infiltration inhibition; alpha d integrin;

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;

KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;

KW atherosclerosis; multiple sclerosis; aschma; psoriasis; Crohn's disease;

KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;

KW rheumatoid arthritis; central nervous system injury; CD11b.

XX Homo sapiens.

OS Homo sapiens.

XX WO200029446-A1.

PN WO200029446-A1.

PD 25-MAY-2000.

XX 25-MAY-2000.

PP 16-NOV-1999; 99WO-US027139.

XX 16-NOV-1999; 99US-00193043.

PR 08-JUL-1999; 99US-00350259.

XX (ICOS-) ICOS CORP.

PA Gallatin MW, Van Der Vieren M;

PI Gallatin MW, Van Der Vieren M;

XX WPI; 2000-387751/33.

DR WPI; 2000-387751/33.

XX WPI; 2000-387751/33.

PT Use of novel anti-alpha integrin d monoclonal antibodies to inhibit

PT macrophage infiltration and reduce inflammation at central nervous system

XX injury sites.

XX Example 5; Fig 1; 270pp; English.

PS Integrins are a class of membrane-associated molecules that participate

CC in cellular adhesion. Integrins are made up of an alpha subunit and a

CC beta subunit. One class of human integrins are restricted to expression

CC in white blood cells and have a common beta2 subunit: the leukocyte

CC integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins

SQ	Sequence 1153 AA;	
Query Match	99.8%; Score 5862; DB 3; Length 1153;	
Best Local Similarity	99.3%; Pred. No. 0;	
Matches 1129; Conservative	7; Mismatches 1; Indels 0; Gaps 0;	
Qy	1 FNLDTENAMTPOENARGFGQSVWOLQGSRRVVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 60	
Db	17 FNLDTENAMTPOENARGFGQSVWOLQGSRRVVVVGAPQEIIVAAANQSGSLYQCDYSTGSCPEI 76	
Qy	61 RLQVPVAVNMSLGLSLAATTSPPQLLACGTPVHTQTCSENTYVKGCLFCLFGSNLRQOPQK 120	
Db	77 RLQVPVAVNMSLGLSLAATTSPPQLLACGTPVHTQTCSENTYVKGCLFCLFGSNLRQOPQK 136	
Qy	121 FPEALRGCPQEDSDIAFLVDGSGIILPHDPRAKEFTSTWMEQLKSKTFLPSLMQYSEEP 180	
Db	137 FPEALRGCPQEDSDIAFLVDGSGIILPHDPRAKEFTSTWMEQLKSKTFLPSLMQYSEEP 196	
Qy	181 RIHFTFEFQNNPNRSLIKPIITQLGRTHGTATGIRKVVRELFTNGARKNAFKILILI 240	
Db	197 RIHFTFEFQNNPNRSLIKPIITQLGRTHGTATGIRKVVRELFTNGARKNAFKILIVI 256	
Qy	241 TDEKFGDPIGYEDVPEADREGVIRYVIGVDADFSEKSRQELNTVASKPPRDHVFQIN 300	
Db	257 TDEKFGDPIGYEDVPEADREGVIRYVIGVDADFSEKSRQELNTIASKPPRDHVFQVN 316	
Qy	301 NFALKTIQNLREKIFAIETGTGSSSSSEHEMSQEGFSAAITNSGPELLSTVGSYDWAG 360	
Db	317 NFALKTIQNLREKIFAIETGTGSSSSSEHEMSQEGFSAAITNSGPELLSTVGSYDWAG 376	
Qy	361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILRNVRQSLVLGAPRYQHIGLVAMFR 420	
Db	377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGVAALILRNVRQSLVLGAPRYQHIGLVAMFR 436	
Qy	421 QNTGMWESNANVKGTOIGYFAGSLCSVDVDSNGSTDVLIGAPHYYEQTREGGQSVYCP 480	
Db	437 QNTGMWESNANVKGTOIGYFAGSLCSVDVDSNGSTDVLIGAPHYYEQTREGGQSVYCP 496	
Qy	481 PRGORARWQCDVLYGQGGQPGWFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAYILF 540	
Db	497 PRGORARWQCDVLYGQGGQPGWFGAALTVLGVDVNGDKLTDVAIGAPGEDNRGAYILF 556	
Qy	541 HGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 600	
Db	557 HGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTGVAQGHVLLRSQ 616	
Qy	601 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRCVLHVQKSTRDRLRREGIQSVVT 660	
Db	617 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRCVLHVQKSTRDRLRREGIQSVVT 676	
Qy	661 YDLALDSGRPHSRVAVNETKNSTRTOVLGLTQTCETLKLQLPNCIEDPVSIVLRNF 720	
Db	677 YDLALDSGRPHSRVAVNETKNSTRTOVLGLTQTCETLKLQLPNCIEDPVSIVLRNF 736	
Qy	721 SLVGTPLSAFGLNRPVLAEDAQRLFTALFPFEKNCNDNTCQDDLSTTFPSMSLDCLVWG 780	
Db	737 SLVGTPLSAFGLNRPVLAEDAQRLFTALFPFEKNCNDNTCQDDLSTTFPSMSLDCLVWG 796	
Qy	781 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 840	
Db	797 GPREFNVTVVRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRSWRLACESASSTEV 856	
Qy	841 SGALKSTSCSINHPFENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTET 900	
Db	857 SGALKSTSCSINHPFENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTET 916	
Qy	901 QLELPVKYAYVMVVTSHGVSTKYNFTASENTSRVMOHYOVSNLGRSPISLVLVPV 960	
Db	917 QLELPVKYAYVMVVTSHGVSTKYNFTASENTSRVMOHYOVSNLGRSPISLVLVPV 976	
Qy	961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1020	
Db	977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCISIAVCQRIQCDIP 1036	

Qy	1021 FFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGQGAFFRSQTET 1080	
Db	1037 FFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGQGAFFRSQTET 1096	
Qy	1081 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKQDMSEGGPPGAEPQ 1137	
Db	1097 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKQYKQDMSEGGPPGAEPQ 1153	
RESULT 3		
AAU80252		
ID	AAU80252 standard; protein; 1153 AA.	
XX	AAU80252;	
XX	DT 15-JUL-2002 (first entry)	
XX	Human integrin 1 alpha-M subunit protein.	
XX	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;	
KW	inflammatory disease; autoimmune disorder; Crohn's disease;	
KW	human immunodeficiency virus; HIV; myocardial infarction;	
KW	Sjorgen's syndrome; rheumatoid arthritis.	
XX	Homo sapiens.	
XX	Key Location/Qualifiers	
PH	Misc-difference 499..500	
FT	/note="Encoded by GGG CAG AGG"	
XX	WC200218583-A2.	
XX	07-MAR-2002.	
XX	31-AUG-2001; 2001WO-US027227.	
XX	01-SEP-2000; 2000US-0229700P.	
XX	(BLOO-) CENT BLOOD RES INC.	
XX	Springer TA, Shimoaka M, Lu C;	
XX	WPI; 2002-382964/41.	
XX	N-PSDB; ABK50046.	
XX	Modified integrin-I or integrin I-like domain polypeptide useful as an	
XX	immunogen to produce antibodies specific to polypeptide, comprises a	
XX	disulfide bond such that polypeptide is stabilized in a desired	
XX	conformation.	
XX	Disclosure; Page 109-112; 112pp; English.	
XX	This invention relates to a modified integrin-I or integrin I-like domain	
XX	polypeptide comprising at least one disulfide bond so that the domain is	
XX	stabilised in a desired conformation. The polypeptide of the invention	
XX	may have antiinflammatory or immunosuppressive activities. The	
XX	polypeptides of the invention have an open conformation and are useful	
XX	as immunogens to produce antibodies that selectively bind to integrin I-	
XX	domain; and for identifying a modulator of integrin activity or of	
XX	interaction of an integrin and a cognate ligand. The polypeptide of the	
XX	invention, or antibodies (preferably anti-LFA-1 antibody) is useful for	
XX	treating or preventing an integrin mediated disorder which is an	
XX	inflammatory or autoimmune disorder in a subject and for inhibiting the	
XX	binding of an integrin to a cognate ligand such as Crohn's disease,	
XX	nephritis; human immunodeficiency virus (HIV), myocardial infarction,	
XX	Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic	
XX	composition comprising the peptide of the invention is useful for	
XX	treating an integrin mediated disorder in a subject. The polypeptides	
XX	and/or active or antigenic fragments are useful as reagents for diagnosis	
XX	of integrin-mediated disorders. The present sequence represents the human	
XX	integrin-1 alpha-M protein subunit used to generate the mutant	
XX	polypeptides of the invention	

[illegible]

SQ Sequence 1153 AA;		
Query Match 99.8%; Score 5862; DB 5; Length 1153;		
Best Local Similarity 99.3%; Pred. No. 0;		
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;		
Qy	1 FNLDTENAMTQENARGGQSVOLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 60	
Db	1 FNLDTENAMTQENARGGQSVOLQGSRRVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI 76	
Qy	61 RLQPVAVNMSLGLSLAATTSPOLLACGTVTHQTCSENTYVYKGLCFGLFSGNLRQPOK 120	
Db	77 RLQPVAVNMSLGLSLAATTSPOLLACGTVTHQTCSENTYVYKGLCFGLFSGNLRQPOK 136	
Qy	121 FPEALRGCPQSDSIAFLVDGSGSIIPHDFRAKEFTSTWEOLEKSKTLFSLMOYSEEF 180	
Db	137 FPEALRGCPQSDSIAFLVDGSGSIIPHDFRMKEFTSTWEOLEKSKTLFSLMOYSEEF 196	
Qy	181 RIHFTFKFQNNPRLSIKEIPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240	
Db	197 RIHFTFKFQNNPRLSLVKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256	
Qy	241 TDGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQELNTVASKPPRDHVFQIN 300	
Db	257 TDGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQELNTIASKPPRDHVFQVN 316	
Qy	301 NFEALKTIQNLREKIFAIEGTQTGSSSPHEMSQEGFAAITSNGPILLSTVGSYDWAG 360	
Db	317 NFEALKTIQNLREKIFAIEGTQTGSSSPHEMSQEGFAAITSNGPILLSTVGSYDWAG 376	
Qy	361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420	
Db	377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436	
Qy	421 QNTGMWESNANVKGITGAYFGASLCSVDVDSNGSTDLVIGAPHYEYQTRGGQVSCPL 480	
Db	437 QNTGMWESNANVKGITGAYFGASLCSVDVDSNGSTDLVIGAPHYEYQTRGGQVSCPL 496	
Qy	481 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLIF 540	
Db	497 PRGORARWQCDVLYGEGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAYLIF 556	
Qy	541 HGTSGSGISPSHSGRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600	
Db	557 HGTSGSGISPSHSGRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616	
Qy	601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVVRVCHVQKSTDRRLREGIOQSVVT 660	
Db	617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVVRVCHVQKSTDRRLREGIOQSVVT 676	
Qy	661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPTVLRNF 720	
Db	677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPTVLRNF 736	
Qy	721 SLVGTPLSAFNLRLPVLAEADAQLFTALFFPEKNCNDNICQDDLSITFFSMLDCLVVG 780	
Db	737 SLVGTPLSAFNLRLPVLAEADAQLFTALFFPEKNCNDNICQDDLSITFFSMLDCLVVG 796	
Qy	781 GPREFNVTVVRNDGDSYRTQVTFPPPLDSLVRKYSTLQNRQSRWRACSSASTEV 840	
Db	797 GPREFNVTVVRNDGDSYRTQVTFPPPLDSLVRKYSTLQNRQSRWRACSSASTEV 856	
Qy	841 SGALKSTSCSINHPPIPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900	
Db	857 SGALKSTSCSINHPPIPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916	
Qy	901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVFLVPV 960	
Db	917 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMOHQYQVSNLQORSLPISLVFLVPV 976	
Qy	961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTIACVQRIQCDIP 1020	
Db	977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELKAPVNCSTIACVQRIQCDIP 1036	
Qy	1021 PFGIOEEFNATLKNLSFDWYIKTSHNHLILIVSTAELFNDVSFTLLPQGAFVRSQET 1080	
Db	1037 PFGIOEEFNATLKNLSFDWYIKTSHNHLILIVSTAELFNDVSFTLLPQGAFVRSQET 1096	
Qy	1081 KVEPFEVPNPPLIVGSSVGSLLLALITAALYKLGFFKROYKDMMSGGPPGAEPQ 1137	
Db	1097 KVEPFEVPNPPLIVGSSVGSLLLALITAALYKLGFFKROYKDMMSGGPPGAEPQ 1153	
RESULT 5		
AAO14428		
ID	AAO14428 standard; protein; 1153 AA.	
XX	AC AAO14428;	
XX	DT 03-MAY-2002 (first entry)	
XX	DE Integrin Mac-1 alpha subunit.	
KW	Mac-1; integrin alpha subunit; variant integrin inserted domain protein;	
KW	open conformation; integrin related inflammatory disorder;	
KW	integrin related immunological disorder; rheumatoid arthritis; ischaemia;	
KW	reperfusion; hypovolemic shock; infarction; cerebral shock;	
KW	viral infection; cancer; gene therapy; vaccine;	
KW	bioactive agent screening.	
XX	Unidentified.	
OS	XX	
PN	WO200204521-A2.	
XX	17-JAN-2002.	
XX	09-JUL-2001; 2001WO-US021805.	
XX	07-JUL-2000; 2000US-0216600P.	
XX	(CALY ) CALIFORNIA INST OF TECHNOLOGY.	
XX	(BLOO-) CENT BLOOD RES.	
XX	Springer T;	
XX	WPI; 2002-148167/19.	
XX	New integrin I domain protein having alteration in at least 2	
PT	noncontiguous regions and exits in an open conformation, useful for	
PT	treating, preventing or suppressing inflammatory or immunological	
PT	disorders.	
XX	Example 1; Fig 1F; 90pp; English.	
PS	The invention comprises structurally biased variant integrin inserted (1)	
CC	domain proteins, wherein the alterations to the protein occur in at least	
CC	two noncontiguous regions. Specifically the variant integrin I domain	
CC	proteins are structurally biased to exist in the open conformation,	
CC	thereby altering the binding ability of the protein. The invention also	
CC	comprises nucleic acids encoding the variant integrin I domain proteins.	
CC	The integrin I domain proteins and nucleic acids are useful for treating,	
CC	preventing or suppressing integrin related inflammatory and immunological	
CC	disorders (e.g. rheumatoid arthritis). The variant integrin I domain	
CC	proteins and nucleic acids can also be used for treating: ischaemia/	
CC	reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral	
CC	infection; and cancer. The variant integrin I domain nucleic acids and	
CC	proteins may be used in gene therapy, as vaccines and to screen for	
CC	bioactive agents. The present amino acid sequence represents the Mac-1	
CC	alpha subunit of integrin	
XX	Sequence 1153 AA;	
SQ	Query Match 99.8%; Score 5862; DB 5; Length 1153;	
	Best Local Similarity 99.3%; Pred. No. 0;	
	Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;	



QY 1 FNLDTENAMTFOENARFGQSVWOLQSGRVVVGPOEIVAAORGSLYQCDYSTGSCBPI 60  
Db |||||  
QY 17 FNLDTENAMTFOENARFGQSVWOLQSGRVVVGPOEIVAAORGSLYQCDYSTGSCBPI 76  
Db |||||  
QY 61 RLQVPVAVNMSLGLSLAATTPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120  
Db |||||  
QY 77 RLQVPVAVNMSLGLSLAATTPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136  
Db |||||  
QY 121 PPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTVMEQLKSKTFLSLMQYSEBF 180  
Db |||||  
QY 137 PPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTVMEQLKSKTFLSLMQYSEBF 196  
Db |||||  
QY 181 RIHTEKEFQNNPRSLIKPITQLGRTHATGIRKVVURELENTNGARKNAFKILILI 240  
Db |||||  
QY 197 RIHTEKEFQNNPRSLIKPITQLGRTHATGIRKVVURELENTNGARKNAFKILIVI 256  
Db |||||  
QY 241 TDGKFGDPLGYEDVIPADREGVIRYVIGDAPFRSEKSEQLMTVASKPRDHVFCPL 300  
Db |||||  
QY 257 TDGKFGDPLGYEDVIPADREGVIRYVIGDAPFRSEKSEQLMTVASKPRDHVFCPL 316  
Db |||||  
QY 301 NFEALKTTQNLREKIFAIETGOTGSSSFHEMSQEGFSAAITSNGLPILTVGSYDWAG 360  
Db |||||  
QY 317 NFEALKTTQNLREKIFAIETGOTGSSSFHEMSQEGFSAAITSNGLPILTVGSYDWAG 376  
Db |||||  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAIILNRNVSOLVGLGAPRYOHIGLVAMFR 420  
Db |||||  
QY 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAIILNRNVSOLVGLGAPRYOHIGLVAMFR 436  
Db |||||  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLI GAPHYEQTGRGGVSVCP 480  
Db |||||  
QY 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVLI GAPHYEQTGRGGVSVCP 496  
Db |||||  
QY 481 PRGQARWCCDAVLVYGOQPMGRGAALTVLVDVNGDKLTVAIGAPGEEDNRGAVYLF 540  
Db |||||  
QY 497 PRGQARWCCDAVLVYGOQPMGRGAALTVLVDVNGDKLTVAIGAPGEEDNRGAVYLF 556  
Db |||||  
QY 541 HGTSGSGISPHSQRISAGSKLSPRIQYFGQSLGGQDLTMGLVLTGVAQGHVLLLRSQ 600  
Db |||||  
QY 557 HGTSGSGISPHSQRISAGSKLSPRIQYFGQSLGGQDLTMGLVLTGVAQGHVLLLRSQ 616  
Db |||||  
QY 601 PVLRYKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660  
Db |||||  
QY 617 PVLRYKAIEMFNPREVARNFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676  
Db |||||  
QY 661 YDLALDSRPHSRAVFNETHKSTRTOVLGHTQTCETILKLOLNCIEDPVSPIVLRNF 720  
Db |||||  
QY 677 YDLALDSRPHSRAVFNETHKSTRTOVLGHTQTCETILKLOLNCIEDPVSPIVLRNF 736  
Db |||||  
QY 721 SLVGTPLSAFNLRPVLAEDAQLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
Db |||||  
QY 737 SLVGTPLSAFNLRPVLAEDAQLFTALPFPEKNCNDNI CODDLSITFSFMSLDCLVVG 796  
Db |||||  
QY 781 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWPLACESASSTEV 840  
Db |||||  
QY 797 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWPLACESASSTEV 856  
Db |||||  
QY 841 SGALKSTCSINHPILFPENSEVTFTNITPDVDSKASIGNKLLKANVTSENMPNKNKEF 900  
Db |||||  
QY 857 SGALKSTCSINHPILFPENSEVTFTNITPDVDSKASIGNKLLKANVTSENMPNKNKEF 916  
Db |||||  
QY 901 QLELPVKAVYVMVTVSHGVSTYKLNFTASENTSRVNHQYQVSNLQSRSLPISLFLVFPV 960  
Db |||||  
QY 917 QLELPVKAVYVMVTVSHGVSTYKLNFTASENTSRVNHQYQVSNLQSRSLPISLFLVFPV 976  
Db |||||  
QY 961 RLNQTVWDRPQVTFSENISSCHYKERLP SHSDFLAELRKAPVNCSTAVCQRIQC DIP 1020  
Db |||||  
QY 977 RLNQTVWDRPQVTFSENISSCHYKERLP SHSDFLAELRKAPVNCSTAVCQRIQC DIP 1036  
Db |||||  
QY 1021 FFIQIEEFNATLKNLSFDWYIKTSHNHLIIVSTAILPNDVSVFTLLPQGAFAVRSQTF 1080  
Db |||||  
QY 1037 FFIQIEEFNATLKNLSFDWYIKTSHNHLIIVSTAILPNDVSVFTLLPQGAFAVRSQTF 1096  
Db |||||  
QY 1081 KVEPPEVPNPLPLIVGSSVGLLALLITALYKLGFFKRYQKDMMSGGPFGAEPQ 1137

Db 1097 KVEPPEVPNPLPLIVGSSVGLLALLITALYKLGFFKRYQKDMMSGGPFGAEPQ 1153  
|||  
RESUT 6  
ADD25615  
ID ADD25615 standard; protein; 1153 AA.  
XX ADD25615;  
XX 15-JAN-2004 (first entry)  
XX Binding domain-immunoglobulin fusion protein-associated protein #85.  
XX Binding domain; immunoglobulin; fusion protein; cytostatic;  
XX antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
XX neuroprotective; hinge region; immunoglobulin heavy chain;  
XX CH2 constant region; CH3 constant region; IgG1;  
XX antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
XX malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
XX rheumatoid arthritis; myasthenia gravis; Grave's disease;  
XX type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
XX Unidentified.  
XX OS  
XX US2003118592-A1.  
XX 26-JUN-2003.  
XX 25-JUL-2002; 2002US-00207655.  
XX 17-JAN-2001; 2001US-0367358P.  
XX 17-JAN-2002; 2002US-00053530.  
XX 03-JUN-2002; 2002US-0385691P.  
XX (GENE-) GENE-CRAFT INC.  
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
XX WPI; 2003-803137/75.  
XX New binding domain-immunoglobulin fusion protein, useful for treating a  
XX subject having or suspected of having a malignant condition or a B-cell  
XX disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
XX Disclosure; SEQ ID NO 176; 157pp; English.  
XX The invention relates to a binding domain-immunoglobulin fusion protein  
XX comprising a binding domain polypeptide that is fused to an  
XX immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
XX CH2 constant region polypeptide that is fused to the hinge region  
XX polypeptide, and an immunoglobulin heavy chain CH3 constant region  
XX polypeptide that is fused to the CH2 constant region polypeptide. The  
XX hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
XX hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
XX region polypeptide, derived from (a) having 3 or more cysteine residues;  
XX where the mutated human IgG1 immunoglobulin hinge region polypeptide  
XX contains 2 cysteine residues, where the first cysteine is not mutated; a  
XX mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
XX (a) having 3 or more cysteine residues, where the mutated human IgG1  
XX immunoglobulin hinge region polypeptide contains no more than one  
XX cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
XX polypeptide, derived from (a) having 3 or more cysteine residues; where  
XX the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
XX no cysteine residues. The binding domain-immunoglobulin fusion protein is  
XX capable of at least one immunological activity comprising antibody  
XX dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
XX binding domain polypeptide is capable of specifically binding to an  
XX antigen. Also included are an isolated polynucleotide encoding the  
XX binding domain-immunoglobulin fusion protein, a recombinant expression  
XX construct comprising the polynucleotide operably linked to a promoter,  
XX a host cell transformed or transfected with a recombinant expression  
XX construct, producing the binding domain-immunoglobulin fusion protein, a

CC pharmaceutical composition comprising the binding domain-immunoglobulin  
 CC fusion protein or polynucleotide and a carrier, and treating a subject  
 CC having or suspected of having a malignant condition or a B-cell disorder.  
 CC The binding domain-immunoglobulin fusion protein is useful for treating a  
 CC subject having or suspected of having a malignant condition or a B-cell  
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple  
 CC sclerosis or autoimmune disease. The present sequence is a binding domain  
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
 CC sequence data for this patent formed part of the printed specification  
 CC and is also available in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not  
 CC identified the sequences in the printed specification by their SEQ ID  
 CC number therefore none of the sequences can be explicitly identified.  
 CC  
 XX  
 SQ Sequence 1153 AA;

Query Match 99.8%; Score 5862; DB 7; Length 1153;  
 Best Local Similarity 99.3%; Pred. No. 0;  
 Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVLOGSRVVGAPQBIIVANQSGSLYQCDYSTGSCPEI 60  
 DB 17 FNLDTENAMTFOENARGFGQSVVLOGSRVVGAPQBIIVANQSGSLYQCDYSTGSCPEI 76  
 QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOPK 120  
 DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCLFLFGSNLRQPOPK 136  
 QY 121 FPALRGCPQEDSDIAFLVDGSGSIIIPHDFRRAKEFISTVMEQLKSKTFLSLMQSSEF 180  
 DB 137 FPALRGCPQEDSDIAFLVDGSGSIIIPHDFRRAKEFISTVMEQLKSKTFLSLMQSSEF 196  
 QY 181 RIHFTKEFQNNPNRSLIKPIQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
 DB 197 RIHFTKEFQNNPNRSLIKPIQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256  
 QY 241 TDEKFGDPLGYEDVPEADREGVIRYVIGVGAFAFSEKSRQELNTVASKPPRDHVFQIN 300  
 DB 257 TDEKFGDPLGYEDVPEADREGVIRYVIGVGAFAFSEKSRQELNTVASKPPRDHVFQIN 316  
 QY 301 NFALKTIQNLREKIFAIBGTQTGSSSSPEHMSQEGFSAAITNSGPLLSTVGSYDWAQ 360  
 DB 317 NFALKTIQNLREKIFAIBGTQTGSSSSPEHMSQEGFSAAITNSGPLLSTVGSYDWAQ 376  
 QY 361 GVFLYTSKEKSTINTRVDSMDNDAYLGVAALIIILNRVQSLVIGAPRYQHIGLVAMFR 420  
 DB 377 GVFLYTSKEKSTINTRVDSMDNDAYLGVAALIIILNRVQSLVIGAPRYQHIGLVAMFR 436  
 QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSCPL 480  
 DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSCPL 496  
 QY 481 PRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
 DB 497 PRQARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 556  
 QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQVFGQSLGGQDLTMDGLVDTLVGAQGHVLLRSQ 600  
 DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQVFGQSLGGQDLTMDGLVDTLVGAQGHVLLRSQ 616  
 QY 601 PVLVRKAIMFENPREVARNVFECDQVVKGEAGEVRVCLVHVKSTDRDLREGQIOSVVT 660  
 DB 617 PVLVRKAIMFENPREVARNVFECDQVVKGEAGEVRVCLVHVKSTDRDLREGQIOSVVT 676  
 QY 661 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTETLKLQPNCIEDPVSPIVLRINF 720  
 DB 677 YDLALDSGRPHSRVAFNETKNSTRQTQVLGLTQTCTETLKLQPNCIEDPVSPIVLRINF 736  
 QY 721 SLVGTPLSARCNLRPVLAEDAQLFTALPFPEKNCNDNITCQDDLSITFSFMSLDCLVVG 780  
 DB 737 SLVGTPLSARCNLRPVLAEDAQLFTALPFPEKNCNDNITCQDDLSITFSFMSLDCLVVG 796

QY 781 GPREENVTIVRNDGDSYRTQVTFEPFPLDLSYRKUSTLONORSORSWRLACESASSTEV 840  
 DB 797 GPREENVTIVRNDGDSYRTQVTFEPFPLDLSYRKUSTLONORSORSWRLACESASSTEV 856  
 QY 841 SGALKSTSCSINHPIFPENSEVTFNITFOVDSKASLGKLLKANVTSENMPRTNKTEF 900  
 DB 857 SGALKSTSCSINHPIFPENSEVTFNITFOVDSKASLGKLLKANVTSENMPRTNKTEF 916  
 QY 901 QLELPVKIAYVMVTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSILPISLVFLVPV 960  
 DB 917 QLELPVKIAYVMVTVSHGVSTKYLNFTASENTSRVMOHQYQVSNLQORSILPISLVFLVPV 976  
 QY 961 RLNQTVIWDROPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIOCDIP 1020  
 DB 977 RLNQTVIWDROPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIOCDIP 1036  
 QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQET 1080  
 DB 1037 FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQET 1096  
 QY 1081 KVEPFVEVPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1137  
 DB 1097 KVEPFVEVPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ 1153  
 RESULT 7  
 AAR04136  
 ID AAR04136 standard; protein; 1153 AA.  
 XX AAR04136;  
 AC AC  
 XX 09-SEP-2004 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 07-SEP-1990 (first entry)  
 XX  
 DE Alpha subunit of Mac-1 leukocyte adhesion receptor.  
 XX  
 KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;  
 KW non-specific defence system; integrin gene superfamily.  
 XX  
 OS Synthetic.  
 PH Key  
 FT Region  
 FT 1..116  
 FT /label= signal\_peptide  
 FT Modified-site 86..88  
 FT /label= putative N-glycosylation site  
 FT Modified-site 240..242  
 FT /label= putative N-glycosylation site  
 FT Modified-site 391..393  
 FT /label= putative N-glycosylation site  
 FT Modified-site 469..471  
 FT /label= putative N-glycosylation site  
 FT Modified-site 693..695  
 FT /label= putative N-glycosylation site  
 FT Modified-site 697..699  
 FT /label= putative N-glycosylation site  
 FT Modified-site 735..737  
 FT /label= putative N-glycosylation site  
 FT Modified-site 802..804  
 FT /label= putative N-glycosylation site  
 FT Modified-site 881..883  
 FT /label= putative N-glycosylation site  
 FT Modified-site 901..903  
 FT /label= putative N-glycosylation site  
 FT Modified-site 912..914  
 FT /label= putative N-glycosylation site  
 FT Modified-site 941..943  
 FT /label= putative N-glycosylation site  
 FT Modified-site 947..949  
 FT /label= putative N-glycosylation site  
 FT Modified-site 979..981  
 FT /label= putative N-glycosylation site

PT Modified-site 994..996  
 /label= putative N-glycosylation site  
 1022..1024  
 PT Modified-site  
 /label= putative N-glycosylation site  
 1045..1047  
 FT Modified-site  
 /label= putative N-glycosylation site  
 1051..1053  
 FT Modified-site  
 /label= putative N-glycosylation site  
 1076..1078  
 FT Modified-site  
 /label= putative N-glycosylation site  
 1106..1134  
 FT Region  
 /label= putative\_transmembrane\_region  
 XX  
 EP364690-A.  
 XX  
 XX  
 PD 25-APR-1990.  
 XX  
 XX 17-AUG-1989; 89EP-00115159.  
 XX  
 XX 23-AUG-1988; 88US-00235353.  
 PR 09-MAR-1989; 89US-00321239.  
 XX  
 XX (DAND ) DANA FARBER CANCER INST INC.  
 PA  
 XX  
 PI Springer TA, Corbi A;  
 XX  
 XX WPI; 1990-125938/17.  
 DR N-PSDB; AAQ04043.  
 DR

XX  
 PT New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating  
 PT inflammation and viral infections, and in diagnosis.  
 XX  
 PS Disclosure; Page ?; 3pp; English.  
 CC  
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.  
 CC recognition of and migration to sites of inflammation. It also attaches  
 CC to cellular substrates as part of this function making it useful in  
 CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene  
 CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25  
 CC -MAR-2003 to correct PA field.)  
 CC  
 CC Revised record issued on 09-SEP-2004 : Correction to feature table key  
 CC and pages  
 CC  
 XX Sequence 1153 AA;

Query Match 99.6%; Score 5852; DB 2; Length 1153;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1128; Conservative 7; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 FNLDTENAMTQENARGFQSVOLQGRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 60  
 DB 17 FNLDTENAMTQENARGFQSVOLQGRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI 76  
 QY 61 RLOVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFGLNLRQOQPK 120  
 DB 77 RLOVPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVKGCLFGLFGLNLRQOQPK 136  
 QY 121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDPRAKBFIITVMQKSKTFLSLMOYSBEF 180  
 DB 137 FPEALRGCPQSDSDIAFLVDGSGSIIPHDPRAKBFIITVMQKSKTFLSLMOYSBEF 196  
 QY 181 RIHFTFKFQNNPNSRLIKPTQLLGRHTHTATGIRKVVRELFNITNGARKNAFKILILI 240  
 DB 197 RIHFTFKFQNNPNSRLIKPTQLLGRHTHTATGIRKVVRELFNITNGARKNAFKILIVI 256  
 QY 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDRHFQIN 300  
 DB 257 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDRHFQIN 316  
 QY 301 NFEALKTIONQREKIFAIEGHTQGTSSSFEHMSQEGFSAAITNGPLLSTVGSYDWAG 360  
 DB 317 NFEALKTIONQREKIFAIEGHTQGTSSSFEHMSQEGFSAAITNGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLVGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 420  
 DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLVGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR 436  
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 480  
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLGAPHYEQTRGGQVSVCP 496  
 QY 481 PRGORARWQCDVAVLYGQGPWGRFGAALTIVLGVNGDKLTDVAIGARGEEDNAGVYLF 540  
 DB 497 PRGORARWQCDVAVLYGQGPWGRFGAALTIVLGVNGDKLTDVAIGARGEEDNAGVYLF 556  
 QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAOCHVLLLRASQ 600  
 DB 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAOCHVLLLRASQ 616  
 QY 601 PVLRVKAIMFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
 DB 617 PVLRVKAIMFNPREVARNVFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
 QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
 DB 677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736  
 QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
 DB 737 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPLDLISYRKVSTLQNRQSRWRLACESASSTEV 840  
 DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPLDLISYRKVSTLQNRQSRWRLACESASSTEV 856  
 QY 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLIKANVTSENWPRNKTEF 900  
 DB 857 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLIKANVTSENWPRNKTEF 916  
 QY 901 QLELPVKYAVVWVTVSHGVSTKYLNFTASENSTRVMQHOYQVSNLQSRSLPSLVFLVFP 960  
 DB 917 QLELPVKYAVVWVTVSHGVSTKYLNFTASENSTRVMQHOYQVSNLQSRSLPSLVFLVFP 976  
 QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDPLAELKAPVNVNCSIAVCQRIQCDIP 1020  
 DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDPLAELKAPVNVNCSIAVCQRIQCDIP 1036  
 QY 1021 PFGIQEEFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVBSQDET 1080  
 DB 1037 PFGIQEEFNATLKGNLSPDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVBSQDET 1096  
 QY 1081 KVEPFEVNPPLTVGVSSVGGLLILALITALYKLGPFKQYKDMWSEGPPGABPQ 1137  
 DB 1097 KVEPFEVNPPLTVGVSSVGGLLILALITALYKLGPFKQYKDMWSEGPPGABPQ 1153  
 RESULT 8  
 ADM99589  
 ID ADM99589 standard; protein; 1152 AA.  
 XX  
 AC ADM99589;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Human integrin alphaM subunit precursor protein.  
 XX  
 KW integrin alpha subunit; beta; antipsoriatic; thrombolytic; anticoagulant;  
 KW osteopathic; cytostatic; immunosuppressive; antiinflammatory;  
 KW neuroprotective; antistickling; immunotherapy; inflammatory;  
 KW autoimmune disorder; thrombosis; cancer; osteoporosis;  
 KW sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;  
 KW alphaM.  
 XX  
 OS Homo sapiens.  
 XX

FH Key Location/Qualifiers  
 FT Misc-difference 965 /note= "Encoded by CCC"  
 XX W02004007530-A2.  
 XX 22-JAN-2004.  
 XX 17-JUL-2003; 2003WO-US022301.  
 XX 17-JUL-2002; 2002US-0396783P.  
 PR 17-JUL-2002; 2002US-0396790P.  
 PR 11-SEP-2002; 2002US-0410135P.  
 XX  
 XX (BLOO-) CENT BLOOD RES INC.  
 XX  
 XX Springer TA, Takagi J;  
 XX WPI; 2004-122877/12.  
 DR N-PSDB; ADM99588.  
 XX  
 XX Novel modified integrin protein having extracellular domains of integrin  
 PT alpha and beta subunits or integrin alpha1 and beta3 subunit, useful for  
 PT treating integrin mediated disorders.  
 XX  
 XX Disclosure; SEQ ID NO 4; 232pp; English.  
 XX  
 XX The invention relates to a novel isolated or recombinant modified  
 CC integrin protein having extracellular domains of integrin alpha and beta  
 CC subunits where one of the subunits has one or more mutations, an altered  
 CC surface feature or an amino acid substitution or internal deletion,  
 CC extracellular domains of the integrin beta subunit that comprise a  
 CC mutation that alters a non-cysteine residue to cysteine or extracellular  
 CC domains of integrin alpha and beta subunits. The polypeptide of the  
 CC invention demonstrates antiproliferative, thrombolytic, anticoagulant,  
 CC osteoprotective, cyclostatic, immunosuppressive, antiinflammatory,  
 CC neuroprotective and antisticking activities and may be useful for  
 CC immunotherapy in order to prevent or treat an integrin-mediated disorder  
 CC such as an inflammatory disorder, an autoimmune disorder, thrombosis,  
 CC cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple  
 CC sclerosis. The current sequence is that of the human integrin alphaM  
 CC subunit precursor protein of the invention.  
 XX  
 XX Sequence 1152 AA;  
 SQ  
 Query Match 99.5%; Score 5846.5; DB 8; Length 1152;  
 Best Local Similarity 99.2%; Pred. No. 0;  
 Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;  
 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 60  
 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVANQSGSLYQCDYSTGSCPEI 76  
 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVKGCLFGLGSLNLRQOPQK 120  
 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTYVKGCLFGLGSLNLRQOPQK 136  
 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFLISTMEQLKSKTILFSLMQYSEEF 180  
 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRRAKEFLISTMEQLKSKTILFSLMQYSEEF 196  
 181 RIHFTFKEFQNNPNRSLKIPITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
 197 RIHFTFKEFQNNPNRSLKIPITQLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256  
 241 TDEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQIN 300  
 257 TDEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNTVASKPPRDHVFQIN 316  
 301 NFALKTIONQLEKIFAISGTQTGSSSSFEHMSQEGFSAAITSNGPILLSVSTGSDWAG 360  
 317 NFALKTIONQLEKIFAISGTQTGSSSSFEHMSQEGFSAAITSNGPILLSVSTGSDWAG 376

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLPAPRYOHIGLVAMFR 420  
 DB 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVGLPAPRYOHIGLVAMFR 436  
 QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTYQTRGGQVSVCP 480  
 DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYTYQTRGGQVSVCP 496  
 QY 481 PRGQRAWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEEEDNRGAVYLF 540  
 DB 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEEEDNRGAVYLF 555  
 QY 541 HGTSGSGISPSHSORIASKLSRLOVFGOSLGGQDLTMDGLVDLTGAGCHVILLRSQ 600  
 DB 556 HGTSGSGISPSHSORIASKLSRLOVFGOSLGGQDLTMDGLVDLTGAGCHVILLRSQ 615  
 QY 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRCVCLHVQKSTRDRLEGGIOQSVVT 660  
 DB 616 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRCVCLHVQKSTRDRLEGGIOQSVVT 675  
 QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVLRINF 720  
 DB 676 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPSPVLRINF 735  
 QY 721 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSTIFSPMSLDCLVVG 780  
 DB 736 SLVGTPLSAFNLRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSTIFSPMSLDCLVVG 795  
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
 DB 796 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 855  
 QY 841 SGALKSTSCSINHPPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900  
 DB 856 SGALKSTSCSINHPPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 915  
 QY 901 QLELPVKYAVYVMTSHGVSTKYLNFNTASNTSRVMOHOYOVNLSGORSIPISLVLVVPV 960  
 DB 916 QLELPVKYAVYVMTSHGVSTKYLNFNTASNTSRVMOHOYOVNLSGORSIPISLVLVVPV 975  
 QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP 1020  
 DB 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWNCISIAVCORIQCDIP 1035  
 QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSOTET 1080  
 DB 1036 FFGIQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFVRSOTET 1095  
 QY 1081 KVBPFEPVNPPLIIVGSSVGLLLALLITAAALYKLGFFKQYKDMMSGEGPGABPQ 1137  
 DB 1096 KVBPFEPVNPPLIIVGSSVGLLLALLITAAALYKLGFFKQYKDMMSGEGPGABPQ 1152  
 RESULT 9  
 ADP12435  
 ID ADP12435 standard; protein; 1152 AA.  
 XX  
 AC ADP12435;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE Protein encoded by mRNA of the invention #45.  
 XX  
 KW transplant rejection; immune system; rheumatoid arthritis; lupus;  
 KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W02004042346-A2.  
 XX  
 PD 21-MAY-2004.  
 XX  
 PF 24-APR-2003; 2003WO-US012946.

XX 24-APR-2002; 2002US-00131831.  
PR 20-DEC-2002; 2002US-00325899.  
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.  
PI Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;  
PI Rosenberg S;  
XX WPI; 2004-400724/37.  
XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
XX pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
PT rejection, in an individual, comprises detecting the expression level of  
PT the genes.  
XX Claim 65; SEQ ID NO 2444; 1762pp; English.  
PS The present invention relates to diagnosing or monitoring transplant  
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual  
CC comprises detecting the expression level of one or more genes. The  
CC methods, system and kits are useful in diagnosing or monitoring  
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic  
CC islet, lung, bone marrow or stem cell transplant rejection,  
CC xenotransplant rejection or mechanical organ replacement rejection, in an  
CC individual. The method is also useful in assessing the immune status of  
CC an individual. The methods are also useful in diagnosing and monitoring  
CC diseases that involve the immune system, e.g. rheumatoid arthritis,  
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
CC viral, bacterial or fungal infection. The present sequence represents a  
CC protein that is encoded by the mRNA of the invention.  
XX Sequence 1152 AA;  
SQ

Query Match 99.4%; Score 5839.5; DB 8; Length 1152;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 1127; Conservative 7; Mismatches 2; Indels 1; Gaps 1;

QY 1 FNLDTENAMTQENARGFQSVVQLQGRVVVVGAPQEI VAAANQSGSLVQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTQENARGFQSVVQLQGRVVVVGAPQEI VAAANQSGSLVQCDYSTGSCPEI 76  
QY 61 RLQVVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRLQQPK 120  
DB 77 RLQVVEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLFSGNLRLQQPK 136  
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKKSKTLFSLMOYSBEF 180  
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFISTVMEQLKKSKTLFSLMOYSBEF 196  
QY 181 RIHFTKEFQNNPRSLIKPITOLLGRTHATGIRKVVREL FNTNGARKNAFKILILJ 240  
DB 197 RIHFTKEFQNNPRSLIKPITOLLGRTHATGIRKVVREL FNTNGARKNAFKILVVI 256  
QY 241 TDGEKFGDPLGYEDVPEADRGVIRYVIGVGDAFRSEKSRQELNNTVASKPPRDHVFQIN 300  
DB 257 TDGEKFGDPLGYEDVPEADRGVIRYVIGVGDAFRSEKSRQELNNTVASKPPRDHVFQIN 316  
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNGPLLSVTGSDVWAG 360  
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNGPLLSVTGSDVWAG 376  
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAIILRNVRQSLVGLGAPYOHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGAAIILRNVRQSLVGLGAPYOHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYEQTRGGQSVCP 480  
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIIGAPHYEQTRGGQSVCP 496  
QY 481 PRGQARWQCDVLYGEOGQPKWRFGAALTIVLDVNGDKLTDVAIGAPGEENRGAIVLF 540  
DB 497 PRG-RARWQCDVLYGEOGQPKWRFGAALTIVLDVNGDKLTDVAIGAPGEENRGAIVLF 555

QY 541 HGTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQHVLRLSQ 600  
DB 556 HGTSGSGISPSHSORIASGKLSPLQYFGQSLGGQDLTMDGLVDLTVGAQHVLRLSQ 615  
QY 601 PVLRVKAIMFENPREVARNFECDNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 660  
DB 616 PVLRVKAIMFENPREVARNFECDNDQVVKGEAGEVRVCLHVQKSTRDRRLREGQIQSVVT 675  
QY 661 YDLALDSGRPHSRVAFNETKNSSTRQTVGLGTCTETKLOLPNCIEDPSPVILRLNF 720  
DB 676 YDLALDSGRPHSRVAFNETKNSSTRQTVGLGTCTETKLOLPNCIEDPSPVILRLNF 735  
QY 721 SLVGTPLSAFGNLRPVLAEADQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
DB 736 SLVGTPLSAFGNLRPVLAEADQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 795  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSLACESASSTEV 840  
DB 796 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWSLACESASSTEV 855  
QY 841 SGALKSTSCSINHIFPENSSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
DB 856 SGALKSTSCSINHIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915  
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHYQVSNLQORSLPISLVFLVPV 960  
DB 916 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHYQVSNLQORSLPISLVFLVPV 975  
QY 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCQIP 1020  
DB 976 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCQIP 1035  
QY 1021 PFGQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTEF 1080  
DB 1036 PFGQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSQTEF 1095  
QY 1081 KVEPFEVNPPLPLTVGSSVGGLLIALLITAAALYKLGFFKQYKDMWSEGGPPGAPPQ 1137  
DB 1096 KVEPFEVNPPLPLTVGSSVGGLLIALLITAAALYKLGFFKQYKDMWSEGGPPGAPPQ 1152

RESULT 10  
ADP44061  
ID ADP44061 standard; protein; 1163 AA.  
XX ADP44061;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Human CD11C protein SEQ ID NO:14.  
XX  
KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;  
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.  
XX Homo sapiens.  
XX WO2004053094-A2.  
XX  
PD 24-JUN-2004.  
XX  
PF 08-DEC-2003; 2003WO-US039208.  
XX  
PR 06-DEC-2002; 2002US-0431522P.  
XX  
PA (PPDP-) PPD DEV LP.  
XX  
PI Dunn SJ, Holzmayer TA;  
XX  
DR WPI; 2004-480928/45.  
DR N-PSDB; ADP44060.  
XX  
PT Identifying an inhibitor of human immunodeficiency virus (HIV) entry into

a human host cell useful for preventing and/or treating HIV infection, by identifying an inhibitor of a cell surface polypeptide such as CXCR-4.  
 Claim 1; SEQ ID NO 14; 133pp; English.

The present invention describes a method (M1) for identifying an inhibitor of HIV entry into a human host cell. (M1) comprises identifying an inhibitor of a cell surface polypeptide selected from CXCR-4 (352 amino acids, SEQ ID NO:8, ADP44055), CCR4 (360 amino acids, SEQ ID NO:10, ADP44057), CCR7 (378 amino acids, SEQ ID NO:12, ADP44059), CD11C (1163 amino acids, SEQ ID NO:14, ADP44061), CD47 (323 amino acids, SEQ ID NO:16, ADP44063), CD68 (354 amino acids, SEQ ID NO:18, ADP44065), CD69 (199 amino acids, SEQ ID NO:20, ADP44067), CD74 (566 amino acids, SEQ ID NO:22, ADP44069), CSF3R (836 amino acids, SEQ ID NO:24, ADP44071), RARA (462 amino acids, SEQ ID NO:26, ADP44073), GABBR1 (578 amino acids, SEQ ID NO:28, ADP44075), P2X1 (150 amino acids, SEQ ID NO:30, ADP44077), HELO1 (299 amino acids, SEQ ID NO:32, ADP44079), GPRK6 (576 amino acids, SEQ ID NO:34, ADP44081), or PK2B (1009 amino acids, SEQ ID NO:36, ADP44083). Also described: (1) an inhibitor of a cell surface polypeptide as described above in a human host cell preventing HIV entry into the human host cell; (2) a pharmaceutical composition comprising an inhibitor of (1) and a carrier; and (3) conferring resistance to HIV infection in an individual, by administering the pharmaceutical composition of (2). An inhibitor of HIV entry has anti-HIV and virucide activities, and can be used as an HIV uptake inhibitor. (M1) is useful for identifying protective compounds that inhibit entry of HIV into cells, useful for the prevention and/or treatment of HIV infection. The present sequence represents human integrin alpha X (CD11C), which is used in the exemplification of the present invention.

Sequence 1163 AA;

Query Match 59.3%; Score 3483; DB 8; Length 1163;  
 Best Local Similarity 61.2%; Pred. No. 1e-281;  
 Matches 691; Conservative 141; Mismatches 291; Indels 6; Gaps 4;

1 FNLDTENAMTFOENARFGQSVVOLQSGRVVVGAPQEIIVANQKSGISYQCDYSGCEPI 60  
 20 FNLDTTELTAFRVDSDAGSDGVQYANSVWVVGAPQKITAANQTGGLYQCYSSTGACEPI 79  
 61 RLQVPEAVNMSLGLSLAATTSPLLACGPTVHTCSENTYVKGFLCFLGNSLRQPOPK 120  
 80 GLQVPEAVNMSLGLSLAATTSPLLACGPTVHTCSENTYVKGFLCFLGNSLRQPOPK 137  
 121 FPEALRCGPEDSDIAFLVDGSGIIIPHFERRAKEFLSTVMEQLKSKTLFSLMOYSEEF 180  
 138 LPVSRQCPQEQDIIIVLIDGSGSISSENATMNFVRAVISQFQPSSTQFSLMQFSNKF 197  
 181 RIHTFKFQNNPNRSLIKPIQTLLGRHTATGIRKVVRELFTNITGARKNAFKILILI 240  
 198 QTHFTFEFRSSNPLSLASVHQLQGTFTATAIQNVVHRLPHASYGARRDAKILIVI 257  
 241 TDGKFGDPLGYEDVPEADREGVIRVIGVGAFAFRSEKROELNTVASPPRDHVPQIN 300  
 258 TDGKFGDPLGYEDVPEADREGVIRVIGVGAFAFRSEKROELNTVASPPRDHVPQIN 317  
 301 NFPAKTIQNLREKIPAIETGOTGSSSEHEMSQEGFSAAITNSGPIILSTVGSYDWAG 360  
 318 DFDALXDIQNLREKIPAIETGOTGSSSEHEMSQEGFSAAITNSGPIILSTVGSYDWAG 377  
 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIILNRRVQSLVGLAPRYOHIGLVAMFR 420  
 378 GAFLYPNMSPPTFINMSQENVMDRDSYLGYSYTELALWKGVSQSLVGLAPRYOHIGLVAMFR 437  
 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSDTLVLI GAPHYEYQTRGGQSVCPIL 480  
 438 QVSRQRMKAEVETGTQIGSYFGASLCSVDVDSNGSDTLVLI GAPHYEYQTRGGQSVCPIL 497  
 481 PRGQARWOCDAVLGYBQOPWRFGAALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
 498 PRGWR-RWQWADVLYGEQHPWRFGAALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
 541 HGTSGSGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMDGLVDTLVGAGQGHVLLRSQ 600

Db 557 HGVLPGISPSHSQRIAGSKLSPRLQVFGQSLGGQDLTMDGLVDTLVGAGQGHVLLRSQ 616  
 Qy 601 PVLKVAIMEFNPREVARNVPECNDQVVKGKAGEVRLHVQKSTRDRRLREGOIQSVVT 660  
 Db 617 PVLWVGVMQFIPAEIPRSFAFECREQVVEQTLVQSNICLYIDKRKNLJGSRDLQSSVT 676  
 Qy 661 YDLALDSGRPHSRVAFNETKNSITRQTVGLTQTCETLKLOLNCIEDPVPVILRLNF 720  
 Db 677 LDLDALDGRISPRATFOETKXRSLSRVRLVGLKAKHCFNLLPSCVEDSVTPITRLNF 736  
 Qy 721 SLVGTPLSAFNGRLPVLAEADAQRLFTALFFPEKNCNDNICODDLSITTFSPMSDCLVWG 780  
 Db 737 TLVGKPLLAFLNRLPMLAADAQRYFTASLFFPEKNCNGADHLCQDNLGJSFSPFGLKSLV 796  
 Qy 781 GPREFNVTVVRNDGEDSYNTQVTFPFLDLSTQNVKYSTLQNRQSRQSRWLACESASSTEV 840  
 Db 797 SNLELNAEVMVNDGEDSYGTTITFHPAGLSRYVAEGOKQQLRSILHLTCDSPVVG-- 854  
 Qy 841 SGALKSTSCSINHPPIRPEENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTEF 900  
 Db 855 SQGTWSTSCSINHLIPRGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPTSKTTF 914  
 Qy 901 QLELPVKYAYVMVTSVGHVSTKYLNFAS-ENTSRVVMQHOVQVSNLQORSIPISLVLVP 959  
 Db 915 QLELPVKYAYTVVSSHEQTKYLNFSSEKESHVAMHYQVNNLQORDLPVGINFWVP 974  
 Qy 960 VRLNQTVWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCORIODI 1019  
 Db 975 VELNQEAVMVDVEVSHFPQNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSILAGCLRFCDV 1034  
 Qy 1020 PFGIOSEFNATLKGNSLSDWYIKTSHNHLIIVSTAEILFENDSVFTLLPGOGAFVRSOTE 1079  
 Db 1035 PFSVGEELDTLKGNSLFGWVQILQKKVSVVSVAEITFTSVYSQLPQGEAFMRAQTT 1094  
 Qy 1080 TKVEPEPEVNPFLIIVGSSVGGLLALLALITAAALYKLGFFKQKQKMMSE 1128  
 Db 1095 TVLEKYKVNHTPLIIVGSSIGGLALLALITAVLYKVGFFKQKQKMMSE 1143  
 RESULT 11  
 ADNO2004  
 ID ADNO2004 standard; protein; 1163 AA.  
 XX AC ADNO2004;  
 XX DT 17-JUN-2004 (first entry)  
 XX DE Human inflammatory bowel disease marker - CD11c protein.  
 KW marker gene; inflammatory bowel disease; IBD; Fc gamma R IIIa;  
 KW Fc gamma R IIb; Mlg; NKG-2; hexokinase; HMW4; REG III; LPAP; Mip-1 (beta);  
 KW L-selectin; EGFL6; IDO; IL-8; CD11c; TLR2; ulcerative colitis;  
 KW Crohn's disease.  
 OS Homo sapiens.  
 XX JP200404065120-A.  
 XX PD 04-MAR-2004.  
 XX PF 07-AUG-2002; 2002JP-00229705.  
 XX PR 07-AUG-2002; 2002JP-00229705.  
 XX PA (SUMI) SUMITOMO SEIVAKU KK.  
 XX DR WPI; 2004-209124/20.  
 XX DR N-PSDB; ADNO1990.  
 PT Novel marker gene of inflammatory bowel disease (IBD) comprising base  
 PT sequence of e.g., Fc gamma R IIIa, Fc gamma R IIb, Mlg, NKG-2,  
 PT hexokinase 3, HMW4, CD11c, TLR2 gene, useful for screening therapeutic





Query Match	59.0%;	Score 3469;	DB 8;	Length 1163;
Best Local Similarity	61.0%;	Pred. No. 1.5e-280;		
Matches 689;	Conservative 141;	Mismatches 293;	Indels 6;	Gaps 4
QY	1	FNLDTENAMTFOENARGFGOSVVLQGSVVVGAPOEIVVAAANORGSLYQCDYSTGSCBPI 60		
		: : :		
Db	20	FNLDTIELTAFRVDSAGFGDSVVQYANSVVGAPOKIITAAQTGLGYQCGYSTGACBPI 79		
QY	61	RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHOTCSENTTVKGLCFGLFGSLNQPOK 120		
		: : :		
Db	80	GLQVPEAVNMSLGLSLASTTSPSOLLACGPVTHHECGRNWLTGLCFLLGPT--QLTOR 137		
QY	121	PFEARLGCPOEDSDIAFLVDGSGSIIPHDPRAKFISTVMBOLKSKTLFSLIMOYSBEF 180		
		: : :		
Db	138	LPVSRQECPRQDQIVFLIDGSGSISSRNFAFMNFRVAVISQFORPSTQPSLMQFSNKF 197		
QY	181	RTHFTPKFQNNPNRSLKPIITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240		
		: : :		
Db	198	QTHFTFEFRRTSNPLUSLASHVHQJQGYTTATATQNVVHRLFHASYGARRDATILIVI 257		
QY	241	TDGKFGDPLGYEDVIPADREGVIRYVIGVDGDAFRSKROELINTVASKPPDRDHVFOIN 300		
		: : :		
Db	258	TDGKKEGSDLYKDVIPWADAAGIIRYALGVGLAFQNRNSWKEINDIAKPSQEHIFKVE 317		
QY	301	NPEALKTTQONLREKIFAIEGHTQSGSSFEHMSQEGFSAAITSGNPLLTSTVGSYDWAG 360		
		: : :		
Db	318	DFDALKDIOQKKEKIFAIEGTETTTSSSFELEMAQEGFSAVFTPDGPVLGAVGVSFTWSG 377		
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAVLGYAAAIILNRVQSLATGAPRYQHIGLVAMFR 420		
		: : :		
Db	378	GAFLYPPNMSPTFINNSQENVDMRUSYLGYSTELALWKGVSQSLVLAGAPRYQHTGRAVIFT 437		
QY	421	QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVLGAPHYYBQTRGGQVSVCP 480		
		: : :		
Db	438	QVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTLVLGAPHYYBQTRGGQVSVCP 497		
QY	481	PRGOBARQCDAVLYGEGQGPWRGPAALTVLGDVNGDKLTDVAIGAFGEEDNRGAVYLF 540		
		: : :		
Db	498	PRGNR--RWCDAVLYGEGHPWRGPAALTVLGDVNGDKLTDVVI GAFGEENRGAVYLF 556		
QY	541	HGTSGSGTSPSHSORIAGSKLSPLOYPGOSISGGQDITMDGLVDLTVGAQGHVILLRSQ 600		
		: : :		
Db	557	HGVLGPSLPSHSORIAGSKLSPLOYPGQALSQGGQDITDGLVDLAVGARGQVLLLR 616		
QY	601	PVLRYKAIMBENPREVARNPECDNDVVVKGEAGEVRVCLHVQKSTRDLREGOIQSVVT 660		
		: : :		
Db	617	PVLWVGYSMQIPAEIIPSAFECREQVVSQETLVQSNICLYIDKRSKNLGSRLDQSSVT 676		
QY	661	YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQCTETLKLQLPNCIEDPVSPTVLRINF 720		
		: : :		
Db	677	LDLALDPGLRSPRATFQETKNRSLSRVRVLGLKAHCENFNLLLPSCVEDSVTPTLRINF 736		
QY	721	SLVGTPLSAFONLRPLVAEDAORLTALFPPEKNCNDNICODDLSTTFPFMSLDCLVVG 780		
		: : :		
Db	737	TLVGKPLLAFLNLRPLMAALAQRTTASLFPPEKNCGDHICQDNLGTSFSPGLKSLVVG 796		
QY	781	GPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSQRWBRLACESASSTEV 840		
		: : :		
Db	797	SNLELNAEVMVWNDGEDSYGTTITFSPAGLSYRYVAEGQKQQLRSLHLTCDSPAGV-- 854		
QY	841	SGALKSTCSINHPHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900		
		: : :		
Db	855	SQGTWSTCSRINHILIFRGAQITPLATFDVSPKAVLGRLLLTANVSSENNTPTKSTTF 914		
QY	901	QLELPVKYAVVMVVTSHGVSTKYILNFITAS--ENTSRVMOHQYOVSNLQORSIPTLSLFLVP 959		
		: : :		
Db	915	QLELPVKYAVTVVSSHEQTKYILNFSEBSEKSHVAMHRYQVNNLQORLPYSINFWVP 974		
QY	960	VLRLQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSTAVCORIQCDI 1019		
		: : :		
Db	975	VELNQEAVMVMDVEVSHQPONSLRCSSEKIAPPASDFLAHQKQVPLDCSITAGCLRFCDV 1034		
QY	1020	PFGIQEEFNATLKNLISFOWYIKTSHNHLILVSTAEILFENDSVFTLLPQCGAFVRSOTE 1079		

clone lambda X47 was isolated from a cDNA library constructed from total RNA extracted from phorbol myristate acetate stimulated HL-60 myelomonocytic cells. The library was screened with oligonucleotide probes based on tryptic peptide fragments of p150.95, the sequence can be attached to appropriate control elements and expressed in prokaryotic and eukaryotic cells. The protein can be used to treat or prevent rhinoviral infection because it interacts with ICM-1 and inhibits cell-virus attachment. It can also be used as an anti-inflammatory agent. See also RAQ06063-4, AAQ06068, AAQ07121-8 and AAQ07152-6. (Updated on 25-MAR-2003 RAQ06063-4, AAQ06068, AAQ07121-8 and AAQ07152-6 to correct PA field.)

Sequence 1163 AA;

Query Match	58.9%;	Score 3460;	DB 2;	Length 1163;
Best Local Similarity	60.9%;	Pred. No. 8.7e-280;		
Matches 687;	Conservative 141;	Mismatches 295;	Indels 6;	Gaps 4;

QY	1	FNLDTENAMTFOENARGFGQSQWOLQCSRVVVGAPQEIIVAAORGSILYQCDYSTGSCBPI	60
Db	20	FNLDTTEITAFRVDSDAGFGDSVQVANSVWVGAPQKITAANQTGGILYQCSYGACBPI	79
QY	61	RLQVPVEAVNMNLSGLSLAAATTPPQLACGPTVHQTCSENTRYVXKGLCFLFGSNLRQDPQK	120
Db	80	GLQVPPPEAVNMNLSGLSLASTTSPQLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR	137
QY	121	PPEALRGCPQEDSDIAFLVDGSGSII PHDFRRAKEFISTVMEQLKSKTLESLMQYSEEF	180
Db	138	LPVSRQECPRQEQDIVFLIDSGSGTSSRNFAFMNFRAVLSQFORPSTQFSLMQFSNKF	197
QY	181	RIHETFEFQNNPNSRLIKETITOLLGRTHTATGIRKVVURELFNITNGARKNAFKILILI	240
Db	198	QTHTEFEFRSTNPLSLASVHQLOGTYTATAIQNVVHRLPHASYGARRDAYKILIVI	25
QY	241	TGDEKFGDPLGYEDVIPADREGVIRYVIGDAPRSEKSRQELMTVASKPPRDHVFOIN	300
Db	258	TGKKKEGSDLYKOVIPMADAAGIIRYALVGCLAFQNRNSWKNELNDIAKSPQSEHIFKVE	317
QY	301	NPEALKTTONQIRKFI PAIEGTQOTGSSSSFEHEMSQEGFSAAITSNGLPSTTVGSGVDAG	360
Db	318	DFDALKDITQNLKEKIFAIEGTETSSSSFELEMAQEGFSAVFTPDGTVLGAVGSGFTWSG	377
QY	361	GVELYTSKEKSTFINMTRVDSMDNDAYILGYAAAIILNRVOSLVLCAPRYOHIGILVAMFR	420
Db	378	GAFLYPPNMSPFTFINNSQENVDMRDSYLGYSTELALWKGVQSVLGLAPRYOHTGKAVFT	437
QY	421	QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYBOTRGQGVSVCP	480
Db	438	QVSRQWRMKA EVTGTQIGSYFGASLCSVDVDTGSTDVLIGAPHYYBOTRGQGVSVCP	497
QY	481	PRGQPARWCDAVLYXGEOQPWGRGAALTVLGDVNGDKLTDAVTCAPGEEDNRCGAVIL	540
Db	498	PRGMW-RWMCDAVLYXGEOQHPWGRFGAALTVLGDVNGDKLTDAVTCAPGEENRCGAVIL	556
QY	541	HGTSGSGTSPSHSQRIAGSKLSPRLQYFQCSLSGGQDLTMDGLNDITVGAQGHVLLLRSQ	600
Db	557	HGVLGSPISPSHSQRIAGSQLSRRLQYFQALSGGGQDLTQDGLVDLAVGARGQVLLLR	616
QY	601	PVLRVKAIIMEFNPREVARNVPCNDQVVKGEAGEVVRVCLHVQKSTRDLRGQTSVVT	660
Db	617	PVLWVGVSQMFTPAETIPRSFAFECREQVWSEQTLVQSNICLYIDKRSKNLLGSRDLQSSVT	676
QY	661	YDLALDSGRPHSPRAVENETKNSTRQTVLGTQTCETLKLQLOPCIEDPVSPVILRLNF	720
Db	677	LDLALDPGLRSPRATFOETKNSLSRVRVLGLKAHCENFNLLLPSCVEDUSVFTITLRLNF	736
QY	721	SLVGTPLSAFGNLRPVIAEDAORLETFALFPFKEKNCNDNICQDLSITFTSFMSLDCLVVG	780
Db	737	TLVGKPLLAFLRNLRLPMLAALQRYFTASLPFKEKNCAGDHCQDNLGTSFSPGLKSLLVG	796
QY	781	GPREFNVTVVRNDGEDSVRTQVTFPPFLDLVSRYKUSTIQNQRSQSRWLA CESASSTEV	840
Db	797	SNLELNAEVMWMDGEDSVGTITTFHSPAGLSYRYVARGQOQLRSLHLTCTDSA--PVA	854

Qy	841	SGALKSTCSINHIPIEPENSEVFNTITFDVDSKASLGNKLLLKANTVTSNNMPTNKTFE	900
Db	855	SQGTWSTSCRINHLIFRGAQITFLATFDVSPKAVLGDRLLLLTANVSSNNTPRTSQTTF	914
Qy	901	OLELPVKYAVYVVTSHGVSTKYLNETAS-ENTSRVMQHOYQVSNLQORSIPISLVLFLVP	959
Db	915	OLELPVKYAVYTVSSHEQFTKYLNESEEEKSHVAMHRYQVNNLQORDLPVSNFWVP	974
Qy	960	VRINQTVIMDRPOVTSSENLSSTCHTKERLPSHSDLELAELKAPVNCSTAVCORLOCDI	1019
Db	975	VELNQAVMMDVSVHPQPSRLSCSQAPPASDFLAHQKPNLVDCSLACGLRFRCDV	1034
Qy	1020	PFEGIOEFENATLKGNSLFDWYIKTSHNHLIIYSTABILPNDSVFTLLPQGFAPVRSQTE	1079
Db	1035	PSFSVQBELDFTLLKGNSLFGVWRQILQKKYSVSVSVABITDTSVYSQLPQGRFMRQOTT	1094
Qy	1080	TKVEPPEVPNPPLPIVGVSSVGGILLALIALIATAALYKLGFFRRQYKDMME	1128
Db	1095	TVLEKVKRNPTPLIVGSSIGALLALIALITAVLYKGVFFRRQYKEMME	1143

## RESULT 14

RESOLUT 14  
AAW65091  
ID AAW65091 standard: 1163 AA.

AC AAW65091:

XX  
NT  
28-SEP-1998 (first entry)XX  
DE Human Beta-integrin CD11c subunit protein;

XX Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
 KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
 KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;  
 KW rheumatoid arthritis.  
 KW

Homo sapiens.

XX  
DN  
1195738533-AXX  
PD 17-MAP-1998XX  
DE 07 JUN-1995. 9519-00485618

XX-107

PR 05-AUG-1994; 94US-00286889.

XX  
PA (ICOS-) ICOS CORP.

PT van Der Vieren M. Gallatin WM:

XX  
DP  
WPT: 1998-206565/18-

XX  
PT Screening assay for modulators of integrin binding - using immobilised or  
XX labelled alpha-4 polymerside useful for a treating type-I diabetes.

Example 5. Fig 1A-D: 106pp: English.

This sequence represents a human beta-integrin CD11c subunit which is used to describe a method for identifying compounds that modulate the interaction of the beta-integrin alpha-d subunit with a binding partner of alpha-d which involves contacting an alpha-d polypeptide with an alpha-d binding partner, one of which is immobilised and the other of which is labelled, in the presence of a test compound, and determining if the compound affects binding between the alpha-d polypeptide and alpha-d binding partner, where the alpha-d polypeptide is alpha-d or its fragment comprising the cytoplasmic, transmembrane or extracellular domain of alpha-d. Compounds that modulate alpha-d binding could be used to treat diseases such as type-1 diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome and rheumatoid arthritis

SQ Sequence 1163 AA;

Query Match 58.6%; Score 3446; DB 2; Length 1163;  
Best Local Similarity 60.8%; Pred. No. 1.3e-278;  
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

```
QY 1 FNLDTENAMTFOENARGGQSYVQLQSGRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI 60
DB 20 FNLDTEELTAFRVDSAGFGDSVVQYANSVWVVGAPQKIIAANQIGLVCQGYSTGACBPI 79
QY 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCFGLFSGNLRLQOPQK 120
DB 80 GLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCFGLFSGNLRLQOPQK 137
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIPTVMELKSKTILSLMOYSEEF 180
DB 138 LPVSRQECPRQEDIVFLIDSGSISSRNFAFMNFVRAVISQFORPTQFSLMOFSNKF 197
QY 181 RIHTFFKFOFNPNRSLIKPITOLGRTHTATGIRKVVRELFNITNGARKNAKILILI 240
DB 198 QTHFTFEFRRTSNPLSLASVHQLQGFTYTATATQNVVHRLFHASYGARRDAIKILLI 257
QY 241 TDGKFGDPLGVEDVIPADREGVIRYVIGVDAPRSEKSGEQLMTVASKPRDRHVQIN 300
DB 258 TDGKFGDPLGVEDVIPADREGVIRYVIGVDAPRSEKSGEQLMTVASKPRDRHVQIN 317
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGILLSTVGSYDWAG 360
DB 318 DFDALKIQLQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAITNSGILLSTVGSYDWAG 377
QY 361 GVFIYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVIGAPRYQIHGLVAMFR 420
DB 378 GAFLYPNMSPTFINMSQENVDMRDSYLGYSTELALMKVQSLVIGAPRYQIHGLVAMFR 437
QY 421 QNTCMWESNANVKTOIGAYFASCLSDVDNSGSDTLVLGAPHYEQTGRGGQVSCPL 480
DB 438 QVSRQWRKAEVIGTQISYFASCLSDVDNSGSDTLVLGAPHYEQTGRGGQVSCPL 497
QY 481 PRQARWQCDVLYGQGOBQWGFAGALTVLGVNGDKLTDVAIGAPGEDNKGAYLIF 540
DB 498 PRGWR-RWCDVLYGQGOBQWGFAGALTVLGVNGDKLTDVAIGAPGEDNKGAYLIF 556
QY 541 HGTSGGSISSHQSRIAGSKLSPLOYFGQSLGQDLTMDGLVLTGVAQGHVLLRSQ 600
DB 557 HGVLPSPISPHSQRIAGSKLSPLOYFGQSLGQDLTMDGLVLTGVAQGHVLLRSQ 616
QY 601 PVLRVKALMEFNPREVARNFECDNVVKGKEAGEVRVCLHVQKSTDRDLREGOIQSVT 660
DB 617 PVLWVGVSQFIPABIPRSAPRECEQVYSEQTLLVQSNICLYIDKRKNLLGSRDLQSVT 676
QY 661 YDLALDSGRPHSAVFNFTKNSRROTQVLGLTQTCETLKLQPLNCTEDPVSPVLRLNF 720
DB 677 LDALAPGLRSLPRAIFQETKNSRVRVLGLKAKCNFNLLPSCVEDSVIPIILRLNF 736
QY 721 SLVGTPLSAFNGRLPVLAEDAQRLLFTALFPFKKNGNDNICQDDLSITFSFMSJDCILVWG 780
DB 737 TLVCKPLIAPNRLRPMALAAQRYFTASLPFKKNGADHICQDLNGLSFSFGPKGLSLVG 796
QY 781 GPREEVTVVRNDGDSYQTVTFPPDLSTKRVKSTLQNRQSQRSLRACSSASTEV 840
DB 797 SNLLENAVMVNDGDSYQTVTFPPDLSTKRVKSTLQNRQSQRSLRACSSASTEV 854
QY 841 SGALKSTCSINHIPIPENSEVFNITFDVDSKASLGNKLLKXANTVSENMPRTNKTET 900
DB 855 SQGTWSTSCINHLIFRGQAQITFLATFDVSPKAVGLDRLLLANVSSNNIPRTSKTIF 914
QY 901 QLELPVKYAVYVVSSEHQFTKYNLFNTAS-ENTSRVMQHQYQVSNLQGRSLPSLFLVP 959
DB 915 QLELPVKYAVYVVSSEHQFTKYNLFNTAS-ENTSRVMQHQYQVSNLQGRSLPSLFLVP 974
QY 960 VRLNQTVIDRPOVTSNLSSTCHTKERLPSHDSFLAELRKAPVNVNCSTAVCQIQCDI 1019
DB 975 VELNQBAVMVDVESHQPNSLRCSSEKTAAPPASDFLAHQKPNVLDSCSIAGCLRFRCDV 1034
```

```
QY 1020 PFGIOEEFNATLKNLSFDWYIKTSNHNHLLIVSTAELFNDSVFTLLFGQGFVRSQTE 1079
DB 1035 PFSFVQEEELDTLKNLSFGWVRQLQKKVSVVVAEIIFDTSVYSQLPGQEAQFMAQTI 1094
QY 1080 TKVEPEFVNPPLPLTVGSSVGLLILLALITAAALYKLGFFKQYKQDMSE 1128
DB 1095 TVLEKYKHNPPLPLTVGSSVGLLILLALITAVLYKVGFFKQYKQEMMBE 1143
```

## RESULT 15

```
AA07361
ID AAB07361 standard; protein; 1163 AA.
XX AAB07361;
AC AAB07361;
XX 17-JAN-2001 (first entry)
XX Human CD11c protein sequence.
XX Human; macrophage infiltration inhibition; alpha_d integrin;
XX leukocyte integrin; Leu-CAM; leukointegrin; immune response;
XX inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;
XX atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;
XX lung inflammation; acute respiratory distress syndrome; Crohn's disease;
XX rheumatoid arthritis; central nervous system injury; CD11c.
XX Homo sapiens.
XX WO200029446-A1.
XX 25-MAY-2000.
XX 16-NOV-1999; 99WO-US027139.
XX 16-NOV-1998; 98US-00193043.
XX 08-JUL-1999; 99US-00350259.
XX (ICOS-) ICOS CORP.
XX Gallatin MW, Van Der Vieren M;
XX WPI; 2000-387751/33.
XX Use of novel anti-alpha integrin d monoclonal antibodies to inhibit
XX macrophage infiltration and reduce inflammation at central nervous system
XX injury sites.
XX Example 5; Fig 1; 270pp; English.
```

Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit CD11c. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AAA60014 and AAB07359). The present sequence has approximately 56% identity to the protein sequence of alpha\_d. The alpha\_d gene and protein may be useful in therapy for diseases linked to aberrant alpha\_d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha\_d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease

Sequence 1163 AA;

Query Match 58.6%; Score 3446; DB 3; Length 1163;  
Best Local Similarity 60.8%; Pred. No. 1.3e-278;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds  
(without alignments)  
4119.157 Million cell updates/sec

Title: US-09-902-481B-5  
Perfect score: 5876  
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKMDMSGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5876	100.0	1137	10	US-09-902-481A-5
2	5862	99.8	1153	9	US-09-350-259-3
3	5862	99.8	1153	10	US-09-902-481A-1
4	5862	99.8	1153	10	US-09-891-943-3
5	5862	99.8	1153	14	US-10-144-259-30
6	5862	99.8	1153	14	US-10-207-655-176
7	5855	99.6	1137	10	US-09-902-481A-6
8	5851	99.6	1137	10	US-09-902-481A-4
9	5846.5	99.5	1152	9	US-09-945-265-4
10	5836	99.3	1137	10	US-09-902-481A-3
11	3469	59.0	1163	14	US-10-116-275-204
12	3446	58.6	1163	9	US-09-350-259-4
13	3446	58.6	1163	10	US-09-891-943-4

14	3411	58.0	1161	9	US-09-350-259-2	Sequence 2, Appli
15	3411	58.0	1161	10	US-09-891-943-2	Sequence 2, Appli
16	3395.5	57.8	1161	9	US-09-350-259-99	Sequence 99, Appl
17	3395.5	57.8	1161	10	US-09-891-943-99	Sequence 99, Appl
18	3232.5	55.0	1161	9	US-09-350-259-55	Sequence 55, Appl
19	3232.5	55.0	1161	10	US-09-891-943-55	Sequence 55, Appl
20	3222	54.8	1161	9	US-09-350-259-53	Sequence 53, Appl
21	3222	54.8	1161	10	US-09-891-943-53	Sequence 53, Appl
22	3211.5	54.7	1151	9	US-09-350-259-37	Sequence 37, Appl
23	3211.5	54.7	1151	10	US-09-891-943-37	Sequence 37, Appl
24	3201	54.5	1155	9	US-09-350-259-46	Sequence 46, Appl
25	3201	54.5	1155	10	US-09-891-943-46	Sequence 46, Appl
26	1848	31.4	369	13	US-10-087-192-1212	Sequence 1212, Ap
27	1557.5	26.5	1170	9	US-09-945-265-2	Sequence 2, Appli
28	1553.5	26.4	1170	15	US-10-261-164-1	Sequence 1, Appli
29	1530	26.0	1223	16	US-10-408-765A-295	Sequence 295, App
30	1359.5	23.1	1086	16	US-10-408-765A-1871	Sequence 1871, Ap
31	1229.5	20.9	494	9	US-09-350-259-103	Sequence 103, App
32	1229.5	20.9	494	10	US-09-891-943-103	Sequence 103, App
33	1153	19.6	1179	14	US-10-177-550-2	Sequence 2, Appli
34	1153	19.6	1179	14	US-10-173-551-2	Sequence 2, Appli
35	1151.5	19.6	413	9	US-09-350-259-101	Sequence 101, App
36	1151.5	19.6	413	10	US-09-891-943-101	Sequence 101, App
37	1103.5	18.8	1151	10	US-09-884-130-103	Sequence 103, App
38	1103.5	18.8	1151	10	US-09-836-353A-103	Sequence 103, App
39	1103.5	18.8	1179	10	US-09-918-715-250	Sequence 250, App
40	1103.5	18.8	1179	17	US-10-474-794-250	Sequence 250, App
41	1101	18.7	1188	14	US-10-291-265-810	Sequence 810, App
42	1097	18.7	1188	14	US-10-291-265-338	Sequence 338, App
43	1093.5	18.6	1189	10	US-09-984-130-35	Sequence 35, Appl
44	1093.5	18.6	1189	10	US-09-836-353A-35	Sequence 35, Appl
45	1093.5	18.6	1189	15	US-10-262-839-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1  
US-09-902-481A-5  
; Sequence 5, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
; US-09-902-481A-5

Query Match	100.0%;	Score 5876;	DB 10;	Length 1137;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1137;	Conservative	0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	FNLDTENAMTFQENARGFGQSVVOLGSRVVGAPQEI	VAANQSGSLYQCDYSTGSCPEI	60
Db	1	FNLDTENAMTFQENARGFGQSVVOLGSRVVGAPQEI	VAANQSGSLYQCDYSTGSCPEI	60
QY	61	RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQT	CSNTYVKGLCFLFGSNLRQOPQK	120
Db	61	RLQVPVEAVNMSLGLSLAATTSPQQLACGPTVHQT	CSNTYVKGLCFLFGSNLRQOPQK	120

```
121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFTSTVMEQLKKSKTLFSLMQYSEEF 180
Db 121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFTSTVMEQLKKSKTLFSLMQYSEEF 180
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILLI 240
Db 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILLI 240
Qy 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHVQFN 300
Db 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHVQFN 300
Qy 301 NFEALKTIONOLREKIFAIEGTQGTSSSFHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
Db 301 NFEALKTIONOLREKIFAIEGTQGTSSSFHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Qy 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
Db 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
Qy 481 PRGQARWQCDAVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 481 PRGQARWQCDAVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Qy 601 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 601 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Qy 661 YDLALDSGRPHRAVFNETHKSTRQTQVLGTLTQTCETLKLQIPNCIEDPVSPIVLRNF 720
Db 661 YDLALDSGRPHRAVFNETHKSTRQTQVLGTLTQTCETLKLQIPNCIEDPVSPIVLRNF 720
Qy 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSPMSLDCLVVG 780
Db 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFPEKNCNDNICODDLSITFSPMSLDCLVVG 780
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWFLACASSTEV 840
Db 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSWFLACASSTEV 840
Qy 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNPRTNKTFF 900
Db 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENNPRTNKTFF 900
Qy 901 QLELPVKYAVMVVTSHGVSSTYKLNFTASENTSRVMQHOYQVSNLQORSPLISLVLFPV 960
Db 901 QLELPVKYAVMVVTSHGVSSTYKLNFTASENTSRVMQHOYQVSNLQORSPLISLVLFPV 960
Qy 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDIP 1020
Db 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDIP 1020
Qy 1021 FPGIOEEFNATLKGNLSPDWYIKTSHNHLIIVSTABILFNDVSFTLLPGGAFVRSQTET 1080
Db 1021 FPGIOEEFNATLKGNLSPDWYIKTSHNHLIIVSTABILFNDVSFTLLPGGAFVRSQTET 1080
Qy 1081 KVEPFPEVNPPLPIVGVSSVGLLILLALITAALYKLGFFKRYQKXMMSEGGPPGAEPO 1137
Db 1081 KVEPFPEVNPPLPIVGVSSVGLLILLALITAALYKLGFFKRYQKXMMSEGGPPGAEPO 1137

Query Match 99.8%; Score 5862; DB 9; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFOENARGFGQSVVQIQGSRVVVGAPQEIIVAAANQSGSLYCCDYSTGSCBPI 60
Db 17 FNLDTENAMTFOENARGFGQSVVQIQGSRVVVGAPQEIIVAAANQSGSLYCCDYSTGSCBPI 76
Qy 61 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCFGLFNSLQQPOK 120
Db 77 RLQVPEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTYVVKGLCFGLFNSLQQPOK 136
Qy 121 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFTSTVMEQLKKSKTLFSLMQYSEEF 180
Db 137 FPEALRGCPQSDSDIAFLVDGSGSIIPHDFRRAKEFTSTVMEQLKKSKTLFSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILLI 240
Db 197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILLI 256
Qy 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHVQFN 300
Db 257 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPFRSEKSRQELNTVASKPPRDHVQFN 316
Qy 301 NFEALKTIONOLREKIFAIEGTQGTSSSFHEMSQEGFSAAITSNGLPSTVGSYDWAG 360
Db 317 NFEALKTIONOLREKIFAIEGTQGTSSSFHEMSQEGFSAAITSNGLPSTVGSYDWAG 376
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436
Qy 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 480
Db 437 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEQTRGGQVSVCP 496
Qy 481 PRGQARWQCDAVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 497 PRGQARWQCDAVLYGEGQPGWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTVGAQGHVLLRSQ 616
Qy 601 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660
Db 617 PVLRYKAIIMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676
```

QY 661 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720  
DB 677 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840  
DB 797 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 900  
DB 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVVMVTSYKYNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV 960  
DB 917 QLELPVKYAVVMVTSYKYNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV 976  
QY 961 RLNQTVIWDROPQVTSSENLSSCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
DB 977 RLNQTVIWDROPQVTSSENLSSCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1036  
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1080  
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1096  
QY 1081 KVEPFEVNPPLIIVGSSVGGLLLALITAALYKLGFFKRYKQKDMWSEGGPPGAEPQ 1137  
DB 1097 KVEPFEVNPPLIIVGSSVGGLLLALITAALYKLGFFKRYKQKDMWSEGGPPGAEPQ 1153

## RESULT 3

US-09-902-481A-1  
; Sequence 1, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (17)-()  
; OTHER INFORMATION:  
US-09-902-481A-1

Query Match 99.8%; Score 5862; DB 10; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTPOENARFGQSVWLQGSRRVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTPOENARFGQSVWLQGSRRVVGAPQEIIVANQRGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPEAVNMNLSGLSLAATTSPPQLIACGTVHTQTCSENTYVYKGLCFPLFGSNLRQOPQK 120  
DB 77 RLQVPEAVNMNLSGLSLAATTSPPQLIACGTVHTQTCSENTYVYKGLCFPLFGSNLRQOPQK 136

## RESULT 4

US-09-891-943-3  
; Sequence 3, Application US/09891943  
; Publication No. US2003007728A1

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKBEFISTVMBQLKSKTLFSLMOYSEBF 180  
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKBEFISTVMBQLKSKTLFSLMOYSEBF 196  
QY 181 RHHTFFKFXFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNTINGARKNAFKLILILI 240  
DB 197 RHHTFFKFXFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNTINGARKNAFKLILVI 256  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNVTASKPPDRHVFQIN 300  
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSRQELNVTASKPPDRHVFQIN 316  
QY 301 NFEALKTQNLREKIFAIEGTQTCSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360  
DB 317 NFEALKTQNLREKIFAIEGTQTCSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLAGAPRYOHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLAGAPRYOHIGLVAMFR 436  
QY 421 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSVCP 480  
DB 437 QNTGWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSVCP 496  
QY 481 PRGORARQCCDAVLYGEOGPWGRFGAALTIVLGDWNGKLTDAIAGAPCEEDNRGAVYLF 540  
DB 497 PRGORARQCCDAVLYGEOGPWGRFGAALTIVLGDWNGKLTDAIAGAPCEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHRSORIGASKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
DB 557 HGTSGSGISPSHRSORIGASKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
QY 601 PVLKVAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLKVAIMEFNPREVARNVFECDNVVKGKEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720  
DB 677 YDLALDSGRPHSRVAVNETKSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALPFPFKNGCNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840  
DB 797 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 900  
DB 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVVMVTSYKYNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV 960  
DB 917 QLELPVKYAVVMVTSYKYNFTASENTSRVMQHQYQVSNLQSRSLPISLVFLVPV 976  
QY 961 RLNQTVIWDROPQVTSSENLSSCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
DB 977 RLNQTVIWDROPQVTSSENLSSCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDIP 1036  
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1080  
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1096  
QY 1081 KVEPFEVNPPLIIVGSSVGGLLLALITAALYKLGFFKRYKQKDMWSEGGPPGAEPQ 1137  
DB 1097 KVEPFEVNPPLIIVGSSVGGLLLALITAALYKLGFFKRYKQKDMWSEGGPPGAEPQ 1153



```
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: NO. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match          99.8%; Score 5862; DB 10; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQOLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQOLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPDPRRAKEFIPTVMEQKKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHPDPRRAKEFIPTVMEQKKSKTLFSLMQYSEEF 196
QY 181 RIHFTKEFQNNPNRSLKIPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RIHFTKEFQNNPNRSLKIPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVITPEADREGVIRYVIGVDGDAFSEKSRQELNIVASKPRDHDVQFN 300
DB 257 TDGEKFGDPLGYEDVITPEADREGVIRYVIGVDGDAFSEKSRQELNIVASKPRDHDVQFN 316
QY 301 NFEALKTIONQLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDWAQ 360
DB 317 NFEALKTIONQLREKIFAIEGTQTGSSSFEHEMSQEGFSAATISNGPLLLSTVGSYDWAQ 376
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAIAIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGVAIAIILNRVQSLVLGAPRYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVSCPL 480
DB 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEQTRGGQSVSCPL 496
QY 481 PRQORARWQCDAYLGYEQGPWGRFGAALTVLGDVNGDKLITDVAIGAPGEEDNRGAYILF 540
DB 497 PRQORARWQCDAYLGYEQGPWGRFGAALTVLGDVNGDKLITDVAIGAPGEEDNRGAYILF 556
QY 541 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGTSGSGISPSHSQRTAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 616
QY 601 PVLRVKAIMFENPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGIOISVVT 660
DB 617 PVLRVKAIMFENPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTEDRLREGIOISVVT 676
QY 661 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVPISVILRLNF 720
```

```
DB 677 YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLPNCIEDPVPISVILRLNF 736
QY 721 SLVGTPELSAFGNLRPVLAEADAQRLFTALFPPEKKNCGNDNICODDLSITFFSFMSLDCLVVG 780
DB 737 SLVGTPELSAFGNLRPVLAEADAQRLFTALFPPEKKNCGNDNICODDLSITFFSFMSLDCLVVG 796
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFFFPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
DB 797 GPREFNVTVTVRNDGEDSYRTQVTFFFPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKYAVVMVYTSHGVSSTKYNLFTASENTRVMQHQYQVSNLQSRSLPISLVFLVPV 960
DB 917 QLELPVKYAVVMVYTSHGVSSTKYNLFTASENTRVMQHQYQVSNLQSRSLPISLVFLVPV 976
QY 961 RINQTVIWRDPQVTFESENLSSTCHTKERLPHSDPFLAELRKAPVNVNCSIAVCQRIQCDIP 1020
DB 977 RINQTVIWRDPQVTFESENLSSTCHTKERLPHSDPFLAELRKAPVNVNCSIAVCQRIQCDIP 1036
QY 1021 PFGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPCOGAFVRSQTEI 1080
DB 1037 PFGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILLFNDVSFTLLPCOGAFVRSQTEI 1096
QY 1081 KVEPPEVNPPLIIVGSSVGGLLALALITAAALYKLGFPKQYKDMWSEGGPPGABEQ 1137
DB 1097 KVEPPEVNPPLIIVGSSVGGLLALALITAAALYKLGFPKQYKDMWSEGGPPGABEQ 1153

RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30

Query Match          99.8%; Score 5862; DB 14; Length 1153;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPQOLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPQOLLACGPTVHQTCSENTYVKGCLFLFGSNLRQOPQK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHPDPRRAKEFIPTVMEQKKSKTLFSLMQYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHPDPRRAKEFIPTVMEQKKSKTLFSLMQYSEEF 196
QY 181 RIHFTKEFQNNPNRSLKIPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
```

```
Db 197 RIHFTFKFONNPNRSLVKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVILGVGDARFSEKSRQELNIVASKPRDRHVFQIN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVILGVGDARFSEKSRQELNIVASKPRDRHVFQIN 316
QY 301 NFEALKTIONQOLREKI FAIEGTQTGSSSSFEHMSQEGFSAATITSGPLLSVGVSDWAG 360
Db 317 NFEALKTIONQOLREKI FAIEGTQTGSSSSFEHMSQEGFSAATITSGPLLSVGVSDWAG 376
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVIGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVIGAPYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGITGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGITGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQVSVCP 496
QY 481 PRGQARWQCDVILYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 540
Db 497 PRGQARWQCDVILYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 556
QY 541 HGTSGGSI SPHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGGSI SPHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
QY 601 PVLKVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
Db 617 PVLKVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRVAVNETKNSTRRTQVGLGTQTCETLKLQPNCIDBPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAVNETKNSTRRTQVGLGTQTCETLKLQPNCIDBPVSPVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDRQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDRQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916
QY 901 QLELPVKAVYVMVTSYHGVSTKYLNTASNTSRVMQHOVQVSNLQORSUPISLVFLVPV 960
Db 917 QLELPVKAVYVMVTSYHGVSTKYLNTASNTSRVMQHOVQVSNLQORSUPISLVFLVPV 976
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCIP 1020
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCIP 1036
QY 1021 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1080
Db 1037 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTE 1096
QY 1081 KVEPPEVPNPLVIGSSVCGLLILALITAALYKLGFQKQYKDMMSGEGPPGABPQ 1137
Db 1097 KVEPPEVPNPLVIGSSVCGLLILALITAALYKLGFQKQYKDMMSGEGPPGABPQ 1153
```

## RESULT 6

```
US-10-207-655-176
; Sequence 176, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
```

```
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 176
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-176
```

## Query Match

```
Best Local Similarity 99.8%; Score 5862; DB 14; Length 1153;
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 60
Db 17 FNLDTENAMTQENARGFGQSVVQLQGSRRVVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPEI 76
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLLIACGPTVHQTCSNTYVVKGLCFPGSNLRSQPOK 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLLIACGPTVHQTCSNTYVVKGLCFPGSNLRSQPOK 136
QY 121 FPEALRGCPQEDSDIAFLVDGSGSII PHDFERRAKEFISTVMEOLKSKTLFSLMOYSBEF 180
Db 137 FPEALRGCPQEDSDIAFLVDGSGSII PHDFERRAKEFISTVMEOLKSKTLFSLMOYSBEF 196
QY 181 RIHFTFKFONNPNRSLVKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
Db 197 RIHFTFKFONNPNRSLVKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVVI 256
QY 241 TDGEKFGDPLGVEDVPEADREGVIRYVILGVGDARFSEKSRQELNIVASKPRDRHVFQIN 300
Db 257 TDGEKFGDPLGVEDVPEADREGVIRYVILGVGDARFSEKSRQELNIVASKPRDRHVFQIN 316
QY 301 NFEALKTIONQOLREKI FAIEGTQTGSSSSFEHMSQEGFSAATITSGPLLSVGVSDWAG 360
Db 317 NFEALKTIONQOLREKI FAIEGTQTGSSSSFEHMSQEGFSAATITSGPLLSVGVSDWAG 376
QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVIGAPYQHIGLVAMFR 420
Db 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVIGAPYQHIGLVAMFR 436
QY 421 QNTGMWESNANVKGITGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGITGAYFGASLCSVDVDSNGSTDLVLIGAPHYEYQTRGGQVSVCP 496
QY 481 PRGQARWQCDVILYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 540
Db 497 PRGQARWQCDVILYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVILF 556
QY 541 HGTSGGSI SPHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 600
Db 557 HGTSGGSI SPHSQRIAGSKLSPRLQYFGOSLSGGQDLTMDGLVDLTGVAQGHVLLRSQ 616
QY 601 PVLKVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 660
Db 617 PVLKVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGQIOSVVT 676
QY 661 YDLALDSGRPHSRVAVNETKNSTRRTQVGLGTQTCETLKLQPNCIDBPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAVNETKNSTRRTQVGLGTQTCETLKLQPNCIDBPVSPVILRLNF 736
QY 721 SLVGTPLSAFGLNRPVLAEDRQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780
Db 737 SLVGTPLSAFGLNRPVLAEDRQRLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 796
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 900
Db 857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPTNKTEF 916
```

QY 901 QLELPVKYAVYVMTVSHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSPLSELVLPV 960  
DB 917 QLELPVKYAVYVMTVSHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSPLSELVLPV 976  
QY 961 RLNTQVIWDRPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020  
DB 977 RLNTQVIWDRPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1036  
QY 1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTAELIENDSVFTLLPGQAGFVRSQTE 1080  
DB 1037 FFGIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTAELIENDSVFTLLPGQAGFVRSQTE 1096  
QY 1081 KVEPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKQYKQVNMSEGGPPGAE 1137  
DB 1097 KVEPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKQYKQVNMSEGGPPGAE 1153

RESULT 7  
US-09-902-481A-6  
; Sequence 6, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimooka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-6

Query Match 99.6%; Score 5855; DB 10; Length 1137;  
Best Local Similarity 99.1%; Pred. No. 0;  
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARFGQSVVQLOGSRVVGAPQEIIVAANQSGSLYQCYSTGSCBPI 60  
DB 1 ENLDTENAMTFQENARFGQSVVQLOGSRVVGAPQEIIVAANQSGSLYQCYSTGSCBPI 60  
QY 61 RLQVPEAVNMSGLSLAATSPOLLACGPTVHQTCSNTYVYKGLCFGLFSGNLROQPK 120  
DB 61 RLQVPEAVNMSGLSLAATSPOLLACGPTVHQTCSNTYVYKGLCFGLFSGNLROQPK 120  
QY 121 FPEALGCCPEDSDIAPLVNDSGSIIPHDERRAKEFTSTVMEQLKSKTLFSLMQYSEEF 180  
DB 121 FPEALGCCPEDSDIAPLVNDSGSIIPHDERRAKEFTSTVMEQLKSKTLFSLMQYSEEF 180  
QY 181 RHIFTFKEFONNPNRSLKPIQTLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
DB 181 RHIFTFKEFONNPNRSLKPIQTLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKRSQELNTVASKPRPDHVFQIN 300  
DB 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKRSQELNTVASKPRPDHVFQIN 300  
QY 301 NFEALKTIQNLREKIFAIRGTOTGSSSPEHEMSQEGFSAATISNGPLLSTVGSYDWAG 360  
DB 301 NFEALKTIQNLREKIFAIRGTOTGSSSPEHEMSQEGFSAATISNGPLLSTVGSYDWAG 360  
QY 361 GVFLYTSKEKSTFFNMTRVDSMDNDAYLGVAIAIILNRVQSLVLGAPRYOHIGLVAMFR 420  
DB 361 GVFLYTSKEKSTFFNMTRVDSMDNDAYLGVAIAIILNRVQSLVLGAPRYOHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQICAYFGASLCSDVDNSGSTDVLIGAPHYYEQTRGGQVSVCP 480  
DB 421 QNTGMWESNANVKGTQICAYFGASLCSDVDNSGSTDVLIGAPHYYEQTRGGQVSVCP 480  
QY 481 PRGQARWQCDVAVLYGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
DB 481 PRGQARWQCDVAVLYGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNRGAVYLF 540  
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSLSGQDITMDGLVDLTVGAOHHVLLLRSQ 600  
DB 541 HGTSGSGISPSHSQRIAGSKLSPLQVFGQSLSGQDITMDGLVDLTVGAOHHVLLLRSQ 600  
QY 601 PVLRVKALMEENPREVARNFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 601 PVLRVKALMEENPREVARNFECDNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLPFFPLDLSYKXVSTLQNRQSRWRLACESASTEV 720  
DB 661 YDLALDSGRPHSRVAFNETKSTRQTQVLPFFPLDLSYKXVSTLQNRQSRWRLACESASTEV 720  
QY 721 SLVGTPLSAFGNLRPVLAEDAQRULTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
DB 721 SLVGTPLSAFGNLRPVLAEDAQRULTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKXVSTLQNRQSRWRLACESASTEV 840  
DB 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYKXVSTLQNRQSRWRLACESASTEV 840  
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
DB 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSPLSELVLPV 960  
DB 901 QLELPVKYAVYVMTVSHGVSTKYLNTASNTSRVMQHOYQVSNLQORSLSPLSELVLPV 960  
QY 961 RLNTQVIWDRPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020  
DB 961 RLNTQVIWDRPQVTFSENLSTCHTKERLPKSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020  
QY 1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTAELIENDSVFTLLPGQAGFVRSQTE 1080  
DB 1021 FFGIQEEFNATLKGNLSFDWYIKTSHNHLIIVSTAELIENDSVFTLLPGQAGFVRSQTE 1080  
QY 1081 KVEPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKQYKQVNMSEGGPPGAE 1137  
DB 1081 KVEPEVENPLPLIVGSSVGGILLALITAAALYKLGFFKQYKQVNMSEGGPPGAE 1137

RESULT 8  
US-09-902-481A-4  
; Sequence 4, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimooka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic

US-09-902-481A-4

Query Match 99.6%; Score 5851; DB 10; Length 1137;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
DB 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPERRAKEPSTVMEOLKSKTFLFSLMOYSEEF 180
DB 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPERRAKEPSTVMEOLKSKTFLFSLMOYSEEF 180

QY 181 RIHFTPKFQONNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 181 RIHFTPKFQONNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 300
DB 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 300

QY 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCP 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCP 480

QY 481 PRGQARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540

QY 541 HGTSSGSLSPHSQRIAGSKLSPLRYQFGSLSGQDITMDGLVLTGAAQGHVLLRSQ 600
DB 541 HGTSSGSLSPHSQRIAGSKLSPLRYQFGSLSGQDITMDGLVLTGAAQGHVLLRSQ 600

QY 601 PVLRVKALIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660
DB 601 PVLRVKALIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660

QY 661 YDLALDSGRPHSRVAFNETKNSTRQTOVLGLTQTCETLKLQPNCTEDPSPVILRLNF 720
DB 661 YDLALDSGRPHSRVAFNETKNSTRQTOVLGLTQTCETLKLQPNCTEDPSPVILRLNF 720

QY 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
DB 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780

QY 781 GPREFNVTVVRNDGDSYKTOVTFPPFLDLSYKVSITLQNRQSRWRLACESASSTEV 840
DB 781 GPREFNVTVVRNDGDSYKTOVTFPPFLDLSYKVSITLQNRQSRWRLACESASSTEV 840

QY 841 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 841 SGALKSTSCSINHPIPFENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900

QY 901 QLELPVKAYVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLQGRSLPI SILVFLPV 960
DB 901 QLELPVKAYVMVTSHGVS TKYLNFTASENTSRVMQHQYQVSNLQGRSLPI SILVFLPV 960

QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCNSIAVCQRIQCDIP 1020
DB 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCNSIAVCQRIQCDIP 1020
```

```
QY 1021 FPGIQEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSOTET 1080
DB 1021 FPGIQEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFFRSOTET 1080

QY 1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPPQ 1137
DB 1081 KVEPFEVPNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEPPQ 1137
```

## RESULT 9

US-09-945-265-4  
; Sequence 4, Application US/09945265  
; Patent No. US20020123614A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy A.  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Iu, Chafen  
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A  
; FILE REFERENCE: CBN-002CP  
; CURRENT APPLICATION NUMBER: US/09/945,265  
; PRIOR FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: US 60/229,700  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 1152  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-945-265-4

Query Match 99.5%; Score 5846.5; DB 9; Length 1152;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1128; Conservative 7; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120
DB 77 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 136

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPERRAKEPSTVMEOLKSKTFLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLVDGSGSIIIPHDPERRAKEPSTVMEOLKSKTFLFSLMOYSEEF 196

QY 181 RIHFTPKFQONNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240
DB 197 RIHFTPKFQONNPNRSLIKPITOLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNTVASKPRDHVFQIN 316

QY 301 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 360
DB 317 NFEALKTIONQREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLSTVGSYDWAG 376

QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCP 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQSVCP 496

QY 481 PRGQARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVLYGEOQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
```

541 HGTSGGISPSHRSORIASGKLSPLRYFGQSLSGQDLTMDGLVDLTGAGQHVLLRSQ 600  
556 HGTSGGISPSHRSORIASGKLSPLRYFGQSLSGQDLTMDGLVDLTGAGQHVLLRSQ 615  
601 PVLRYKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
616 PVLRYKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675  
661 YDLALDSGRPHSRVAFNETKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720  
676 YDLALDSGRPHSRVAFNETKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 735  
721 SLVGTPLSAFNLRLPVLAEADAQLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
736 SLVGTPLSAFNLRLPVLAEADAQLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 795  
781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSWRLACESASSTEV 840  
796 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSWRLACESASSTEV 855  
841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 900  
856 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 915  
901 QLELPVKYAVYVMVTSHGVSITKYNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV 960  
916 QLELPVKYAVYVMVTSHGVSITKYNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV 975  
961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020  
976 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1035  
1021 FFGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080  
1036 FFGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1095  
1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1137  
1096 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1152

RESULT 10  
US-09-902-481A-3  
; Sequence 3, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimooka, Motomu  
; APPLICANT: Shifman, Julia  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-3

Query Match 99.3%; Score 5836; DB 10; Length 1137;  
Best Local Similarity 98.9%; Pred. No. 0;  
Matches 1124; Conservative 7; Mismatches 6; Indels 0; Gaps 0;  
QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVVGAPQEIIVANQORSLYQCDSYTGSCBPI 60  
DB 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVVGAPQEIIVANQORSLYQCDSYTGSCBPI 60

QY 61 RLQVPVBAVNMSLGLSLAATTPPQLLACQPTVHTQCSNTYVKGCLFCLFGSNLRQPOK 120  
DB 61 RLQVPVBAVNMSLGLSLAATTPPQLLACQPTVHTQCSNTYVKGCLFCLFGSNLRQPOK 120  
QY 121 FPEALRGCPQSDSDIAHLVDGSGSIIIPHDPRRAKEFISTVMEOLKSKTLFSLMOYSBEF 180  
DB 121 FPEALRGCPQSDSDIAHLVDGSGSIIIPHDPRRAKEFISTVMEOLKSKTLFSLMOYSBEF 180  
QY 181 RIHETFEFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
DB 181 RIHETFEFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELMTVASKPRDHVFQIN 300  
DB 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSKQELMTVASKPRDHVFQIN 300  
QY 301 NFEALKTQNLREKIPALIEGTOTSSSSFEHMSQEGFSAAITNGPILLSVGSYDWA 360  
DB 301 NFEALKTQNLREKIPALIEGTOTSSSSFEHMSQEGFSAAITNGPILLSVGSYDWA 360  
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420  
DB 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGYAAAIILNRVQSLVLGAPRYQHIGLVAMPR 420  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCP 480  
DB 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEQTRGGQSVCP 480  
QY 481 PRGQARWOCDAVLGEQGPWRGGAALTVDLGVNKGDLTDVAIGAPDEENRGAVYLF 540  
DB 481 PRGQARWOCDAVLGEQGPWRGGAALTVDLGVNKGDLTDVAIGAPDEENRGAVYLF 540  
QY 541 HGTSGSGISPSHRSORIASGKLSPLRYFGQSLSGQDLTMDGLVDLTGAGQHVLLRSQ 600  
DB 541 HGTSGSGISPSHRSORIASGKLSPLRYFGQSLSGQDLTMDGLVDLTGAGQHVLLRSQ 600  
QY 601 PVLRYKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 601 PVLRYKAIMENPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
QY 661 YDLALDSGRPHSRVAFNETKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720  
DB 661 YDLALDSGRPHSRVAFNETKNSRTRQTQVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720  
QY 721 SLVGTPLSAFNLRLPVLAEADAQLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
DB 721 SLVGTPLSAFNLRLPVLAEADAQLFTALFPPEKNCNDNI CODDLSITFSFMSLDCLVVG 780  
QY 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSWRLACESASSTEV 840  
DB 781 GPREFNVTVVRNDGEDSVRTQVTFPPDLDSYRKVSTLQNSORSWRLACESASSTEV 840  
QY 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 900  
DB 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 900  
QY 901 QLELPVKYAVYVMVTSHGVSITKYNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV 960  
DB 901 QLELPVKYAVYVMVTSHGVSITKYNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV 960  
QY 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020  
DB 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCISIAVCQRIQCDIP 1020  
QY 1021 FFGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080  
DB 1021 FFGIOEBFNATLKGNSLFDWYIKTSHNHLIIVSTAEILFNDVSFTLLPGQAFVRSQTEP 1080  
QY 1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1137  
DB 1081 KVEPFEVNPPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMMSGGPPGAEPO 1137

RESULT 11  
US-10-116-275-204  
; Sequence 204, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 204  
; LENGTH: 1163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-275-204

Query Match 59.0%; Score 3469; DB 14; Length 1163;  
Best Local Similarity 61.0%; Pred. No. 2e-285;  
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGGQSVVQGGVGVVGAPOEIVAAQORSLYQCDYSTGSCPEI 60  
Db 20 FNLDTTEALTAFRVDSAGFGDSVVQVANSVWVWVGAPOKITAANQTGGLYQCGYSTGACPEI 79

QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQCSTENTVVKGLCFGLFGSLRQOQOK 120  
Db 80 GLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQCSTENTVVKGLCFGLFGSLRQOQOK 137

QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEQLKSKTFLSLMOYSEF 180  
Db 138 LPVSRQECRQEQDIFVLIDGSGSISSRNFMVFAVAVISQFORPTQSLMOFSNKF 197

QY 181 RIHETFEFQNNPNRSLIKIPITOLLGTHATGIRKRVRELFNITNGARKNAFKILILI 240  
Db 198 QTHETFEFRTSNPLSLASVHQLQGYTATATQNVVHLEFHASYGARDATKILIVI 257

QY 241 TDGKFGDPLCYEDVPEADREGVIRYVIGVDFAFRSEKSRQELNTVASKPRDHVFQIN 300  
Db 258 TDGKKGDSLDYKVIWADDAAGIIRVAIGVGLAFQNNRNWKLNDIASKPSQEHIFKVE 317

QY 301 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFSAATISNGPILLSVGSYDWAG 360  
Db 318 DFDALKDIONLKEKIFAIEGTETTSSTSSFELEMAQEGFSAVFPTDGPVLGAVGSFTWSG 377

QY 361 GVELYTSKEKSTFNMTVRVDSMDNDAVLGYAAAIILNRVQSLVGLAPRYOHIGLVAMFR 420  
Db 378 GAFLYPPNMGPTFINMSQENVDSDYSLGISTELALWKVQSLVGLAPRYOHTKAVIFT 437

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQVSCPL 480  
Db 438 QVSEQWEMKAEVTGTLGISTFGASLCSVDVDTGSTDVLVIGAPHYYEOTRGQVSCPL 497

QY 481 PRGORARQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 498 PRGWR-RWWCDVLYGEGHPWGRFGAALTVLGDVNGDKLTDVVGAPGEEHNRGAVYLF 556

QY 541 HGTSGSISPSHSQRIAGSKLSPQLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
Db 557 HGVLGPISEPSHSQRIAGSKLSPQLQYFGQALSGQDLTQDGLVDLAVGARGQVLLRTR 616

QY 601 PVLVRKAIMFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLRREGQIQSVVT 660  
Db 617 PVLVWVGSMQFIPEAFPRSAFECREQVSEQTIVQSNICILYIDKRKNLIGSRDLQSSVT 676

RESULT 12  
US-09-350-259-4  
; Sequence 4, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vliet, Monica  
; TITLE OF INVENTION: No. US20020062008A1 Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-4

Query Match 58.6%; Score 3446; DB 9; Length 1163;  
Best Local Similarity 60.8%; Pred. No. 1.8e-283;  
Matches 686; Conservative 147; Mismatches 290; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGGQSVVQGGVGVVGAPOEIVAAQORSLYQCDYSTGSCPEI 60  
Db 20 FNLDTTEALTAFRVDSAGFGDSVVQVANSVWVWVGAPOKITAANQTGGLYQCGYSTGACPEI 79

QY 61 RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQCSTENTVVKGLCFGLFGSLRQOQOK 120  
Db 80 GLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQCSTENTVVKGLCFGLFGSLRQOQOK 137







Db 677 LDALAPGLSRAIPQETKNRSLSRVRLGLKAHCENENLLPSCVEDSVIILRLNF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAORLFTALPFEKNCNGDNICQDLSITFSFMSLDCLVVG 780  
Db 737 TLVKGELLAFAFRNRPMLAALQRYFTASLPFEKNCNGADHCQDNLGISFSFGLKSLVG 796  
QY 781 GRPEFNVTVVRNDGSDSYRTQVTFPPPLDLSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
Db 797 SNLELNAEVMVWMDGSDSYRTQVTFPPHAGLSYRYVAEGQKQGLASLHLC--CSAPVG 854  
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900  
Db 855 SGTWSTSCRIHNLIFRGAQITFLATFDVSPKAVGLDRLLLTANVSSENNTPTSKTIF 914  
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFAS-ENTSRVMOHQYOVSNLQSRSLPISLVFLVP 959  
Db 915 QLELPVKYAVYVMTSHGVSTKYLNFSESEKSHVAMHRYOVNNGQDLVPSINFWVP 974  
QY 960 VRLNQTVIWRDPQVTFSENLSSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCQRIQCDI 1019  
Db 975 VELNQEAVMMDEVESHQPNPSLRCSSEKIAPPASDFLAHIQKNPVLDCSIACCLFRCDV 1034  
QY 1020 PFGIOEEFNATLKNLFDWIKTSHNHLIVSTAEILFNDSTVTLPLPGQAFVRSQTE 1079  
Db 1035 PSFSVQEEELDFTLKGNLSFGWVRQILQKKVSVSVAEIIFTDTSVYQLPGQEAFAWQTI 1094  
QY 1080 TKVEPFEPENPLPIVGVSSVGLLILALITAAALYKLGFFKQYKDMWSE 1128  
Db 1095 TVLEKYKVHNPILVGVSSIGLLILALITATLYKLGFFKQYKEMBLE 1143

## RESULT 14

US-09-350-259-2  
; Sequence 2, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20020062008A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350.259  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193.043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173.497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286.889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362.652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943.363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1161  
; TYPE: PRT  
; ORGANISM: Homo sapiens

US-09-350-259-2

Query Match 58.0%; Score 3411; DB 9; Length 1161;  
Best Local Similarity 59.4%; Pred. No. 1.7e-280;  
Matches 671; Conservative 166; Mismatches 284; Indels 8; Gaps 6;

QY 1 ENLDTENAMTFQENARGFQGVVQLQGSRRVVGAPQETVAANQSGLYQCDYSTGSCPEI 60  
Db 17 FNDLVEEPTIFQEDAGFGQGVVQFGGSLRVVGLVWAPLEVAANQTRGLYDCAATGWCQPI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTPSPQLLACGPTVHQCSTENTYKGLCFGLGSLNRQPOK 120  
Db 77 PLHIREAVNMSLGLSLAATTPSPQLLACGPTLHRCVGENSYKSGSLGSLRSW-BIIQT 135

## RESULT 15

US-09-891-943-2  
; Sequence 2, Application US/09891943  
; Publication No. US20030077278A1

GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20030077278A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/891,943  
; CURRENT FILING DATE: 2001-06-26  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1161  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-09-891-943-2

Query Match 58.0%; Score 3411; DB 10; Length 1161;

Best Local Similarity 59.4%; Pred. No. 1.7e-280; Indels 8; Gaps 6;  
Matches 671; Conservative 166; Mismatches 284

QY	1	FNLDTENAMTFOENARGFQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYQCDYSTGSCBPI	60
DB	17	FNLDBEETIFQEDAGGFGQSVVQFGSRLVVGAPLEVVAAANTGRLYDCAATGMCQPI	76
QY	61	RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHOTCSNTVYVKGICFLFGSNLQQPOK	120
DB	77	PLHIRPEAVNMSLGLTAASTNGSRLACGPTLHRCVGENSYSGKSCLLGSRW-EITQT	135
QY	121	FEALRGCPQEDSIAFLVDGSGSIIPHDFRAKEFIITVMEQLKKSKTLFSLMOYSBEF	180
DB	136	VDPATPECPHQEMDIVFLIDGSGSIDQDNFQNMGEVQAVMGQFEGTDLFALMQYSNLL	195
QY	181	RIHFTFKFQNNPNPRLIKPITQLLGRTHRTATGIRKVVRELFNIINGARKNAFKILILI	240
DB	196	KIHFTFTQRTSPSQSLVDPIVLQKGLTFTATGILTVTVTQLFHHKNGARKSAKILIVI	255
QY	241	TDGKFGDPLGYEDVITPEADREGVIYIVGVGDADFSEKSRQELINTVASKPPRDHVQIN	300
DB	256	TDGQKYKDPLEYSDVITPOAEKAGIIRYAIYGVGHAFQGTARQELNTISSAPPODHVFKVD	315
QY	301	NFEALKTIONLREKIPALEGTGTGSSSPFHEHMSQEGFSAALTSNGPLLSITVGSYDWAG	360
DB	316	NFAALGSIQKQLEKIYAVEGTQSRASSSPQHMSQEGFSTALTMGDLFLGAVGSFWSG	375
QY	361	GVFLYTSKEKSTFINNTRVDSMDNDAYLVYAAAIILNRRVQSVILVGLAPRYOHIGLVAMFR	420
DB	376	GAFLYPPNMSPTFINNMQENVDMKDSYLGYSSTELALWKGVQNLVGLAPRYOHTGKAVIFT	435
QY	421	QNTGMHESNANVGTQIGAYFGALCSVDVDSNGSITDLVLI GAPHYYEQTGCGVSVCP	480
DB	436	QVSRQWRKKAETVGTQIGSYFGALCSVDVDSNGSITDLIILIGAPHYYEQTGCGVSVCP	495
QY	481	PRGORARWQCDVILYGOGQWGRFGAALTVLGDNVNGDKLTDVAICAPGEEDNRGAYLF	540
DB	496	PRGORVQWQCDVILYGOGQWGRFGAALTVLGDNVNGDKLTDVAICAPGEEDNRGAYLF	555
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGSLGGQDLTMDGLVDLTVGAGQHVLILRSQ	600
DB	556	HGASESGISPSHSQRIAGSKLSPRLQYFGALSGGQDLTQDGLMDLAVGARGQVILLRSL	615
QY	601	PVLKRVKALMENPREVARNYFECNDQVVKGEAGEVCLHVQKSTRDLRREGQIQSVVT	660
DB	616	PVLKRVKALMENPREVARNYFECNDQVVKGEAGEVCLHVQKSTRDLRREGQIQSVVT	673
QY	661	YDLALDSGRPHSRVAVFNETKSTRRTQTQVLGLTQTCETLKLQLPNCIEDPVPVILRLNF	720

Search completed: November 9, 2004, 12:46:58  
Job time : 102.5 secs

Db	674	FDLALDEGRILTSRAIFNETKNPTLTRKKTGLGHCETLKLKLLPDCVEDVVSPIILLHNF	733
QY	721	SLVGTPLSAPGNLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG	780
DB	734	SLVREPIPSQNLPRVLAVSQDLFTASLFPFEKNCQDGLCEGDLGVTLSFSGIQTITVG	793
QY	781	GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYKVVSTLQNRQSRWRWLACESASSTEV	840
DB	794	SSLELNVITVWNAGEDSYGTIVSVLYVPAGLSHRVSGAQKQPHQSALRLACETV-PTED	852
QY	841	SGALKSTSCSINHIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF	900
DB	853	EG-LRSRCSYNHPIFHEGSGTTFIVTVDVSYKATLGDRLMDLRKASSENKASSSKATP	911
QY	901	QLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVP	959
DB	912	QLELPVKYAVVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLISLVFLVP	971
QY	960	VRNLQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAAPVNVCSIAVCQRIQCDI	1019
DB	972	VLLNGVAVMDVMEAPSQSL-PCVSEKPPQHSDFLTQISRSPMLDCSIADCIQFRCDV	1029
QY	1020	PFGLQEEFNATLKNLSFDWVKTSHNHLIIVSTAEILFNDSVFTLLPGQCAFVRSQTE	1079
DB	1030	PSFSVQEEELDTLKNLSFGWVRETLOKKVLSVSAEITFTSVYSQLPQGEAFMRAQME	1089
QY	1080	TKVEPFEVPNPPLITVGSVGGILLALITAAALYKLGFFKQYKQKDMSE	1128
DB	1090	MVLEDEVYNALPIINGSSVGAALLLALITATLYKLGFFKRYKEMLED	1138

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds  
(without alignments)  
3278.416 Million cell updates/sec

Title: US-09-902-481B-5

Perfect score: 5876

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/1/iaa/PTUS\_COMB.pep: \*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5862	99.8	1153	1	US-08-173-497-3
2	5862	99.8	1153	1	US-08-286-889-3
3	5862	99.8	1153	1	US-08-485-618-3
4	5862	99.8	1153	1	US-08-362-652-3
5	5862	99.8	1153	2	US-08-605-672-3
6	5862	99.8	1153	2	US-08-482-293A-3
7	5862	99.8	1153	2	US-08-943-363-3
8	5862	99.8	1153	3	US-09-193-043-3
9	5862	99.8	1153	4	US-09-688-307A-3
10	5862	99.8	1153	4	US-09-350-259-3
11	5831.5	99.2	1152	2	US-08-476-062A-43
12	5831.5	99.2	1152	5	PCT-US96-01314-43
13	5831.5	99.2	1152	6	5424399-2
14	3469	59.0	1163	2	US-08-476-062A-44
15	3469	59.0	1163	5	PCT-US96-01314-44
16	3446	58.6	1163	1	US-08-173-497-4
17	3446	58.6	1163	1	US-08-286-889-4
18	3446	58.6	1163	1	US-08-485-618-4
19	3446	58.6	1163	1	US-08-362-652-4
20	3446	58.6	1163	2	US-08-605-672-4
21	3446	58.6	1163	2	US-08-482-293A-4
22	3446	58.6	1163	2	US-08-943-363-4
23	3446	58.6	1163	3	US-09-193-043-4
24	3446	58.6	1163	4	US-09-688-307A-4
25	3446	58.6	1161	1	US-09-350-259-4
26	3411	58.0	1161	4	US-08-173-497-2
27	3411	58.0	1161	1	US-08-286-889-2

28	3411	58.0	1161	1	US-08-485-618-2	Sequence 2, Appli
29	3411	58.0	1161	1	US-08-362-652-2	Sequence 2, Appli
30	3411	58.0	1161	2	US-08-605-672-2	Sequence 2, Appli
31	3411	58.0	1161	2	US-08-482-293A-2	Sequence 2, Appli
32	3411	58.0	1161	2	US-08-943-363-2	Sequence 2, Appli
33	3411	58.0	1161	3	US-09-193-043-2	Sequence 2, Appli
34	3411	58.0	1161	4	US-09-688-307A-2	Sequence 2, Appli
35	3411	58.0	1161	4	US-09-350-259-2	Sequence 99, Appl
36	3395.5	57.8	1161	1	US-08-485-618-99	Sequence 99, Appl
37	3395.5	57.8	1161	2	US-08-605-672-99	Sequence 99, Appl
38	3395.5	57.8	1161	2	US-08-482-293A-99	Sequence 99, Appl
39	3395.5	57.8	1161	2	US-08-943-363-99	Sequence 99, Appl
40	3395.5	57.8	1161	3	US-09-193-043-99	Sequence 99, Appl
41	3395.5	57.8	1161	4	US-09-688-307A-99	Sequence 99, Appl
42	3395.5	57.8	1161	4	US-09-350-259-99	Sequence 55, Appl
43	3232.5	55.0	1161	3	US-09-193-043-55	Sequence 55, Appl
44	3232.5	55.0	1161	4	US-09-688-307A-55	Sequence 55, Appl
45	3232.5	55.0	1161	4	US-09-350-259-55	Sequence 55, Appl

#### ALIGNMENTS

RESULT 1  
US-08-173-497-3  
; Sequence 3, Application US/08173497  
; Patent No. 5437958  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van Der Vieren, Monica  
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha  
; TITLE OF INVENTION: Subunit  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 S. Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/173,497  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5437958and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/31363  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-173-497-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPQEIIVAAHQGSLYQCDYSTGSCPI 60  
DB 17 FNLDTENAMTFQENARGFGQSVVQLQGRVVGAPQEIIVAAHQGSLYQCDYSTGSCPI 76

RESULT 2  
US-08-286-889-3  
; Sequence 3, Application US/08286889  
; Patent No. 5470953  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Mich  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: P38,659  
; REFERENCE/DOCKET NUMBER: 27866/32168  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-286-889-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQIIVAAQNRGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQIIVAAQNRGSLYQCDYSTGSCPEI 76  
Qy 61 RLQVPVEAVNMSGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFGLSGNLRQOPQK 120  
Db 77 RLQVPVEAVNMSGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFGLSGNLRQOPQK 136  
Qy 121 PPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEQKKSKTFLSLMQYSEEF 180  
Db 137 PPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEQKKSKTFLSLMQYSEEF 196  
Qy 181 RIHFTTFEFQNNPRSLIKPITQLGRTHATGIRKVVRELFINITGARKNAFKILILI 240  
Db 197 RIHFTTFEFQNNPRSLIKPITQLGRTHATGIRKVVRELFINITGARKNAFKILIVI 256  
Qy 241 TDGSKFGDPLGYEDVPEADREGVIRYVIGDAPRSEKSKQELNTVASKPPRDHVQFN 300  
Db 257 TDGSKFGDPLGYEDVPEADREGVIRYVIGDAPRSEKSKQELNTVASKPPRDHVQFN 316  
Qy 301 NFEALKTIONQLREKIPIAEGTQTGSSSSFEHMSQEGFSAAITNGPILLSVGVSYDWAG 360  
Db 317 NFEALKTIONQLREKIPIAEGTQTGSSSSFEHMSQEGFSAAITNGPILLSVGVSYDWAG 376

Qy 61 RLQVPVEAVNMSGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFGLSGNLRQOPQK 120  
Db 77 RLQVPVEAVNMSGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFGLSGNLRQOPQK 136  
Qy 121 PPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEQKKSKTFLSLMQYSEEF 180  
Db 137 PPEALRGCPQEDSDIAFLVDSGSIIPHDPRRAKEFTSTVMEQKKSKTFLSLMQYSEEF 196  
Qy 181 RIHFTTFEFQNNPRSLIKPITQLGRTHATGIRKVVRELFINITGARKNAFKILILI 240  
Db 197 RIHFTTFEFQNNPRSLIKPITQLGRTHATGIRKVVRELFINITGARKNAFKILIVI 256  
Qy 241 TDGSKFGDPLGYEDVPEADREGVIRYVIGDAPRSEKSKQELNTVASKPPRDHVQFN 300  
Db 257 TDGSKFGDPLGYEDVPEADREGVIRYVIGDAPRSEKSKQELNTVASKPPRDHVQFN 316  
Qy 301 NFEALKTIONQLREKIPIAEGTQTGSSSSFEHMSQEGFSAAITNGPILLSVGVSYDWAG 360  
Db 317 NFEALKTIONQLREKIPIAEGTQTGSSSSFEHMSQEGFSAAITNGPILLSVGVSYDWAG 376  
Qy 361 GVFLYTSKESKFNNTRVDSMDNDAYLGAAAILRNVRQSLVGLGAPRYQHIGLVAMPR 420  
Db 377 GVFLYTSKESKFNNTRVDSMDNDAYLGAAAILRNVRQSLVGLGAPRYQHIGLVAMPR 436  
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 496  
Qy 481 PRGORARWOCDAVLYGEQCPWGRFGAALTIVLDVNGDKLTDAICAPGEEDNRGAVILF 540  
Db 497 PRGORARWOCDAVLYGEQCPWGRFGAALTIVLDVNGDKLTDAICAPGEEDNRGAVILF 556  
Qy 541 HGTSGSGISPSHSORTAGSKLSPLOVFGOSLGGQDLTMDGLVDLTVGAQGHVILLRSQ 600  
Db 557 HGTSGSGISPSHSORTAGSKLSPLOVFGOSLGGQDLTMDGLVDLTVGAQGHVILLRSQ 616  
Qy 601 PVLVRKALMEFNPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDLREGIQISVVT 660  
Db 617 PVLVRKALMEFNPREVARNVFECDQVWKGKEAGEVRVCLHVQKSTRDLREGIQISVVT 676  
Qy 661 YDLALDSGRPHSAVNETKSTRQTVLGLTCTETKLPNCIEDPVSIVLRLNF 720  
Db 677 YDLALDSGRPHSAVNETKSTRQTVLGLTCTETKLPNCIEDPVSIVLRLNF 736  
Qy 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALFPFEKNCNDNICQDDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVSRKVKSTLQNRQSRWRLACESASTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVSRKVKSTLQNRQSRWRLACESASTEV 856  
Qy 841 SGALKSTSCSINHPIFPENSEVTNIIPDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTNIIPDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
Qy 901 QLELPVKYAVYVNTSHGVSTKYNFTASENTSRVQHQYVSNLQGRSLPISLVFLVPV 960  
Db 917 QLELPVKYAVYVNTSHGVSTKYNFTASENTSRVQHQYVSNLQGRSLPISLVFLVPV 976  
Qy 961 RLNQTIVWDRPQVTFSENLSSTCHTKERLPKPSHDFLAELRKPVNCISIAVCQRIQCDIP 1020  
Db 977 RLNQTIVWDRPQVTFSENLSSTCHTKERLPKPSHDFLAELRKPVNCISIAVCQRIQCDIP 1036  
Qy 1021 PFGIOEEFNATLKGNSFDWYIKTSHNLLIVSTAEILFNDVSFTLLPQOGAFVRSQTEI 1080  
Db 1037 PFGIOEEFNATLKGNSFDWYIKTSHNLLIVSTAEILFNDVSFTLLPQOGAFVRSQTEI 1096  
Qy 1081 KVEPEVEPNPLPLIVSSVGGILLALITAAIYKLGFFKQYKQMMSEGGPPGABPQ 1137  
Db 1097 KVEPEVEPNPLPLIVSSVGGILLALITAAIYKLGFFKQYKQMMSEGGPPGABPQ 1153

QY 361 GVFLYTSKESKSTINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMR 420  
DB 377 GVFLYTSKESKSTINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMR 436  
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQGVSVCP 480  
DB 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQGVSVCP 496  
QY 481 PRQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540  
DB 497 PRQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 600  
DB 557 HGTSGSGISPSHSQRIAGSKLSRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 616  
QY 601 PVLRVKAIEMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRLRGQIQSVVT 660  
DB 617 PVLRVKAIEMFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDRLRGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVFNFKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 720  
DB 677 YDLALDSGRPHSRVFNFKNSRTRQTVLGLTQTCETLKLQLPNCIEDPVPVILRLNF 736  
QY 721 SLVGTPLSAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLCLVVG 780  
DB 737 SLVGTPLSAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICQDDLSITFSFMSLCLVVG 796  
QY 781 GPRFNVTVTVRNDGEDSYRQVTFPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
DB 797 GPRFNVTVTVRNDGEDSYRQVTFPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVSKASLGKLLKANVTSENMPRTNKTEF 900  
DB 857 SGALKSTSCSINHPIFPENSEVTNITFDVSKASLGKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVYMWVTSKYNLFTASENTSRVMOHQYVSNLQGRSLPISLVFLVPV 960  
DB 917 QLELPVKYAVYMWVTSKYNLFTASENTSRVMOHQYVSNLQGRSLPISLVFLVPV 976  
QY 961 RLNTQTVLWDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
DB 977 RLNTQTVLWDRPQVTFSENLSSTCHTKERLPKSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1036  
QY 1021 FFGIQEEFNATLKNLSFDMWIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTE 1080  
DB 1037 FFGIQEEFNATLKNLSFDMWIKTSHNHLIVSTAEILFNDVSFTLLPGQAGFVRSQTE 1096  
QY 1081 KVEPFEVNPPLPIVSGVGLLLALITAAALYKLGFFKRYKQKDMMSGEGPPGAEPPQ 1137  
DB 1097 KVEPFEVNPPLPIVSGVGLLLALITAAALYKLGFFKRYKQKDMMSGEGPPGAEPPQ 1153

## RESULT 3

US-08-485-618-3  
; Sequence 3, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,618  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32797  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-485-618-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLQSGSRVVGVVGAPOEIVAAANQSGSLVQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFOENARFGQSVVQLQSGSRVVGVVGAPOEIVAAANQSGSLVQCDYSTGSCPEI 76  
QY 61 RLQVPEAVNMVSLGLSLAATTSPPQLACGPTVHTCSENTYVKGCLPFGSNLRQOPQK 120  
DB 77 RLQVPEAVNMVSLGLSLAATTSPPQLACGPTVHTCSENTYVKGCLPFGSNLRQOPQK 136  
QY 121 FPEALRGCPQEDSDIAFLIDSGSGIIIPDFERRAKEFISTVMEOLKSKTLPSLMQYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDSGSGIIIPDFERRAKEFISTVMEOLKSKTLPSLMQYSEEF 196  
QY 181 RIHFTKFEQNNPNRSLIKPTITQLGRTHATGIRKVVRELFNITNGARKNAFKILILI 240  
DB 197 RIHFTKFEQNNPNRSLIKPTITQLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVVGCDAPRSEKSRQELNTVASKPPRDHVPQIN 300  
DB 257 TDGEKFGDPLGYEDVPEADREGVIRVVGCDAPRSEKSRQELNTVASKPPRDHVPQIN 316  
QY 301 NFEALKTIQNQLREKIFAIEGTQSGSSFEHEMSQEGFSAATTSNGPILLSTVGSYDWAG 360  
DB 317 NFEALKTIQNQLREKIFAIEGTQSGSSFEHEMSQEGFSAATTSNGPILLSTVGSYDWAG 376  
QY 361 GVFLYTSKESKSTINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMR 420  
DB 377 GVFLYTSKESKSTINMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMR 436  
QY 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQGVSVCP 480  
DB 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGQGVSVCP 496  
QY 481 PRQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540  
DB 497 PRQARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSRLQYFGQSLSGGQDLTMDGLVDLTGGAQGHVLLRSQ 600

557 HGTSGSGISPSHQSRIAGSLSPRLQYFGQSLGGDLTMDGLVDLTGAGHVLRLRSQ 616  
601 PVLRVKALMEPNPREVARNPECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 660  
617 PVLRVKALMEPNPREVARNPECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 676  
661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQPNCTEDPVSPIVLRNF 720  
677 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQPNCTEDPVSPIVLRNF 736  
721 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
737 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796  
781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPRTNKTEF 900  
857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSENNMPRTNKTEF 916  
901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVFLVPV 960  
917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVMOHQYQVSNLQORSPLISLVFLVPV 976  
961 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPVNVNCSIAVCORIQCDIP 1020  
977 RLNQTVIWRDPQVTFSENLSSCTCHTKERLPSHDSFLAELRKAPVNVNCSIAVCORIQCDIP 1036  
1021 PFGIOEEFNATLKNLSPOWYIKTSINHLLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1080  
1037 PFGIOEEFNATLKNLSPOWYIKTSINHLLIIVSTAEILFNDVSFTLLPGOGAFVRSQTEF 1096  
1081 KVEPFVPRNPLPLIVGSSVGGLLLALITAAALYKLGFFRKQYKDMSEGGPPGAEPO 1137  
1097 KVEPFVPRNPLPLIVGSSVGGLLLALITAAALYKLGFFRKQYKDMSEGGPPGAEPO 1153

RESULT 4  
US-08-362-652-3  
; Sequence 3, Application US/08362652  
; Patent No. 5766850  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 93  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/362,652  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.

REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32391  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-362-652-3

Query Match 99.8%; Score 5862; DB 1; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAAQNRGSLYQCDYSTGSCBPI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPOLIACGPTVHOTCSENTVVKGLCFGLFSGNLQOPOK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPOLIACGPTVHOTCSENTVVKGLCFGLFSGNLQOPOK 136  
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFFRAKEFISTVMEOLKKSKTLFSLMOYSBEF 180  
Db 137 FPEALRGCPQEDSDIAFLIDSGSIIPHDFFRAKEFISTVMEOLKKSKTLFSLMOYSBEF 196  
QY 181 RIHFTFKFQNNPNRSLIKPITQLGRTHATGIRKVVRELFNTNGARKNAFKLLIL 240  
Db 197 RIHFTFKFQNNPNRSLIKPITQLGRTHATGIRKVVRELFNTNGARKNAFKLLIL 256  
QY 241 TDGEKFGDPLGYEDYVPEADREGVIRYVIGVDGAPSEKSRQELNNTVASKPPRDHVFQIN 300  
Db 257 TDGEKFGDPLGYEDYVPEADREGVIRYVIGVDGAPSEKSRQELNNTVASKPPRDHVFQIN 316  
QY 301 NFEALKTIONQREKIFAIEGTQTSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 360  
Db 317 NFEALKTIONQREKIFAIEGTQTSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINNTRVDSMDNDAYLGVAAILLRNRVQSLVLCAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVSCPL 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQSVSCPL 496  
QY 481 PRGORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYILF 540  
Db 497 PRGORARWOCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYILF 556  
QY 541 HGTSGSGISPSHQSRIAGSLSPRLQYFGQSLGGDLTMDGLVDLTGAGHVLRLRSQ 600  
Db 557 HGTSGSGISPSHQSRIAGSLSPRLQYFGQSLGGDLTMDGLVDLTGAGHVLRLRSQ 616  
QY 601 PVLRVKALMEPNPREVARNPECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 660  
Db 617 PVLRVKALMEPNPREVARNPECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQPNCTEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRVAFNETKSTRQTVLGLTQTCETLKLQPNCTEDPVSPIVLRNF 736  
QY 721 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840

Db 797 GPREFNVTVVRNDGSDSVRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPTNKTEF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMQHQYQVSNLQORSILPISLFLVPV 960  
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMQHQYQVSNLQORSILPISLFLVPV 976  
QY 961 RLNQTVWDRPQVTFSENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020  
Db 977 RLNQTVWDRPQVTFSENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036  
QY 1021 PFGIOEFNATLKNLSPDWYIKTSHNHLLIYSTAEILFNDVSFTLLPGOGAFVRSQTEF 1080  
Db 1037 PFGIOEFNATLKNLSPDWYIKTSHNHLLIYSTAEILFNDVSFTLLPGOGAFVRSQTEF 1096  
QY 1081 KVEPPEVNPLPLIVGSSVGGILLALLALITAAALYKLGFFKROYKDMWSEGGPFCAEPQ 1137  
Db 1097 KVEPPEVNPLPLIVGSSVGGILLALLALITAAALYKLGFFKROYKDMWSEGGPFCAEPQ 1153

## RESULT 5

US-08-605-672-3  
; Sequence 3, Application US/08605672  
; Patent No. 5817515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; FILING DATE: US/08/605,672  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-605-672-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENLDTENAMTFQENARFGQSVVQLQSGSRVVGAPQEIIVAANQRGSLYQCDYTGSCPEI 60  
Db 17 FNLDTENAMTFQENARFGQSVVQLQSGSRVVGAPQEIIVAANQRGSLYQCDYTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHTQCSNTYVKGICFLFGSNLRQPOK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPPLLACGPTVHTQCSNTYVKGICFLFGSNLRQPOK 136  
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIISTVMEQLKKSKTLFSLMQYSEEF 180  
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAKEFIISTVMEQLKKSKTLFSLMQYSEEF 196  
QY 181 RHFTFKCFQONNPNRSLKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVI 240  
Db 197 RHFTFKCFQONNPNRSLKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILVI 256  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNNTVASKPRDRHWFQIN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGDAPFRSEKSRQELNNTVASKPRDRHWFQIN 316  
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSGNPLLSVTGSDWAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITSGNPLLSVTGSDWAG 376  
QY 361 GVELYTSKEKSTFINMTFVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 420  
Db 377 GVELYTSKEKSTFINMTFVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYEGASLCSVDVDSNGSTDLVLIGAPHYEOTEQGGVSCVPL 480  
Db 437 QNTGMWESNANVKGTQIGAYEGASLCSVDVDSNGSTDLVLIGAPHYEOTEQGGVSCVPL 496  
QY 481 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNNGAVYLF 540  
Db 497 PRGQARWQCDVLYGEGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNNGAVYLF 556  
QY 541 HGTSGSGISPSHSORIASGLSPLQYFGQSLSGQDLTMDGLVDLTVGAGHVLLLRQ 600  
Db 557 HGTSGSGISPSHSORIASGLSPLQYFGQSLSGQDLTMDGLVDLTVGAGHVLLLRQ 616  
QY 601 PVLKVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTDRLRREGOIOSVVT 660  
Db 617 PVLKVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTDRLRREGOIOSVVT 676  
QY 661 YDLALDSGRPHSRAVFNETHNSTRRTQVLGLTQTCETLKLQFNCIEDPVSPILRLNF 720  
Db 677 YDLALDSGRPHSRAVFNETHNSTRRTQVLGLTQTCETLKLQFNCIEDPVSPILRLNF 736  
QY 721 SLVGTPLSAFNLRPVLAEDAQLFTALPPEKNCNDNICQDDLSITFSFMSLDCLVWG 780  
Db 737 SLVGTPLSAFNLRPVLAEDAQLFTALPPEKNCNDNICQDDLSITFSFMSLDCLVWG 796  
QY 781 GPREFNVTVVRNDGSDSVRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGSDSVRTQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANKVTSNNMPTNKTEF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMQHQYQVSNLQORSILPISLFLVPV 960  
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMQHQYQVSNLQORSILPISLFLVPV 976  
QY 961 RLNQTVWDRPQVTFSENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1020  
Db 977 RLNQTVWDRPQVTFSENLSSSTCHTKERLPSSHDFLAELRKAPVNCSTAVCQRIQCDIP 1036



QY 1021 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEI 1080  
Db 1037 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEI 1096  
QY 1081 KVEPFEPVNPPLVGVSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137  
Db 1097 KVEPFEPVNPPLVGVSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

RESULT 6  
US-08-482-293A-3  
; Sequence 3, Application US/08482293A  
; Patent No. 5831029  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,293A  
; FILING DATE:  
; CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-482-293A-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;  
Best local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;  
QY 1 FNLDTENAMTPOENARGFGQSVQLQGSRRVVGAPQIIVANQRGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTPOENARGFGQSVQLQGSRRVVGAPQIIVANQRGSLYQCDYSTGSCPEI 76

QY 61 RLQPVAVNMSLGLSLAATTSPPOLLACGFTVHQTCSNTYVKGICFLFGSNLRQOPQK 120  
Db 77 RLQPVAVNMSLGLSLAATTSPPOLLACGFTVHQTCSNTYVKGICFLFGSNLRQOPQK 136  
QY 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAXEFISTVMEQLKSKTLFSLMQYSEEF 180

Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDFRAXEFISTVMEQLKSKTLFSLMQYSEEF 196  
QY 181 RIHFTFRFQNNPRSLKIPITOLLGRTHATGIRKVVRELFINITNGARKNAFKILILI 240  
Db 197 RIHFTFRFQNNPRSLKIPITOLLGRTHATGIRKVVRELFINITNGARKNAFKILVVI 256  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVIVGVDAPFRSEKSEQLNTVASKPRPHVPOIN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRVIVGVDAPFRSEKSEQLNTIASKEPRPHVPOIN 316  
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSSPHEMSQSGFSAITNSGILLTSTVSYDWAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSPHEMSQSGFSAITNSGILLTSTVSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSCPL 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSCPL 496  
QY 481 PRGQARWQCDVAVLYGEOQPGWRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVYLF 540  
Db 497 PRGQARWQCDVAVLYGEOQPGWRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVLTVAQAGHVLLLRQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGGQDLTMDGLVLTVAQAGHVLLLRQ 616  
QY 601 PVLRVKAIEMEENPREVARNVFCNDQVVKGEAGVVRVCLHVQKSTRDLRGQIQSVVT 660  
Db 617 PVLRVKAIEMEENPREVARNVFCNDQVVKGEAGVVRVCLHVQKSTRDLRGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVFNETKNSRQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 720  
Db 677 YDLALDSGRPHSRVFNETKNSRQTVLGTQTCETLKLQLPNCIEDPVPVILRLNF 736  
QY 721 SLVGTPLSAFAGNLRPVLAEDAORLFTALFPFKNCGNDNICODDLSITFSFMSLCLVVG 780  
Db 737 SLVGTPLSAFAGNLRPVLAEDAORLFTALFPFKNCGNDNICODDLSITFSFMSLCLVVG 796  
QY 781 GPREENVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLONQORSQSWLACESASSTEV 840  
Db 797 GPREENVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLONQORSQSWLACESASSTEV 856  
QY 841 SGALKSTCSINHPIFPENSEVTNLTFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTCSINHPIFPENSEVTNLTFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVYVMTVSHGVSTKYINFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 960  
Db 917 QLELPVKYAVYVMTVSHGVSTKYINFTASENTSRVMQHQYQVSNLQORSPLISLVFLVPV 976  
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPDSDFLAELKAPVNCIAVCORICDIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPDSDFLAELKAPVNCIAVCORICDIP 1036  
QY 1021 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEI 1080  
Db 1037 FFGIOEFNATLKNLSFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTEI 1096  
QY 1081 KVEPFEPVNPPLVGVSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137  
Db 1097 KVEPFEPVNPPLVGVSSVGLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

RESULT 7  
US-08-943-363-3  
; Sequence 3, Application US/08943363  
; Patent No. 5837478  
; GENERAL INFORMATION:

APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943,363  
FILING DATE: 5-AUG-1994  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: William Jk., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-943-363-3

Query Match 99.8%; Score 5862; DB 2; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVANORGSLYQCDYSTGSCPEI 76  
QY 61 RLOQVPEAVNMSLGLSLAATTSPPQLIACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQ 120  
DB 77 RLOQVPEAVNMSLGLSLAATTSPPQLIACGPTVHQTCSNTYVKGCLFLFGSNLRQOPQ 136  
QY 121 FPEALRGCPQEDSDIAFLVDGSGIIPHPDRAKEPISITVMEQLKSKTLFSLMQYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLVDGSGIIPHPDRAKEPISITVMEQLKSKTLFSLMQYSEEF 196  
QY 181 RIHFTKFEQNNPNRSLKIPITQLLGRHTATGIRKVVRELFNITNGARKNAFKILIT 240  
DB 197 RIHFTKFEQNNPNRSLKIPITQLLGRHTATGIRKVVRELFNITNGARKNAFKILIT 256  
QY 241 TQGEKFGDPLGHDVTPEDADREGVIRYVIGVDAPFSEKSRQELNIVASKPRDHFQIN 300  
DB 257 TQGEKFGDPLGHDVTPEDADREGVIRYVIGVDAPFSEKSRQELNIVASKPRDHFQIN 316  
QY 301 NFEALKTIONQREKIFATEGTQTGSSSFEHMSQEGFSAITSNGPLSTVGSYDMAG 360  
DB 317 NFEALKTIONQREKIFATEGTQTGSSSFEHMSQEGFSAITSNGPLSTVGSYDMAG 376

QY 361 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPYQHIGLVAMPR 420  
DB 377 GVFLYTSKSKSTFINNTRVDSMDNDAYLGAAAAIILNRVQSLVGLGAPYQHIGLVAMPR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASICSVDVDNSGSTDVLIGAPHYYEOTRGQSVCP 480  
DB 437 QNTGMWESNANVKGTQIGAYFGASICSVDVDNSGSTDVLIGAPHYYEOTRGQSVCP 496  
QY 481 PRGQARWQCCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 540  
DB 497 PRGQARWQCCDAVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVLYF 556  
QY 541 HGTSGSIGSPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTVAQGHVLLLSQ 600  
DB 557 HGTSGSIGSPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLTVAQGHVLLLSQ 616  
QY 601 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660  
DB 617 PVLRVKALMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 720  
DB 677 YDLALDSGRPHSRVAFNETKSTRQTQVLGLTQTCETLKLQLPNCIEDPVPVLRNLF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYKSVSTLQNRQSRWRLACESASSTEV 840  
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPDLISYKSVSTLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 900  
DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 916  
QY 901 QLELPVKYAVVMVTSHGVTKYLNFTASENTSRVMQHQYQVSNLQORSIPISLVFLVPV 960  
DB 917 QLELPVKYAVVMVTSHGVTKYLNFTASENTSRVMQHQYQVSNLQORSIPISLVFLVPV 976  
QY 961 RLNQTVIWDPRPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCQRIQDIP 1020  
DB 977 RLNQTVIWDPRPOVTFSENLSSTCHTKERLPSHSDFLAELKAPVWNCISIAVCQRIQDIP 1036  
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDVSFTLLPGOGAFVRSOTET 1080  
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDVSFTLLPGOGAFVRSOTET 1096  
QY 1081 KVEPPEVPNPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMSEGGPPGAEPQ 1137  
DB 1097 KVEPPEVPNPLPIVGVSSVGLLLALITAALYKLGFFKQYKDMSEGGPPGAEPQ 1153

RESULT 8  
US-09-193-043-3  
; Sequence 3, Application US/09193043  
; Patent No. 6251395  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6251395e1 Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/193,043  
; CURRENT FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03

Db 917 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVMOHQVSNLGRSLPISLVFLVPV 976  
Qy 961 RLNQTVINDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIP 1020  
Db 977 RLNQTVINDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVWVNCISIAVCQRIQCDIP 1036  
Qy 1021 PFGIOEEFNATLKGNSLFDWYIKTSHNELLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1080  
Db 1037 PFGIOEEFNATLKGNSLFDWYIKTSHNELLIVSTAEILFNDVSFTLLPGQAFVRSQTET 1096  
Qy 1081 KVEPFEVNPPLIVGSSVGLLALLITLALITLALYKLGFFKQYKDMSEGGPPGAEPQ 1137  
Db 1097 KVEPFEVNPPLIVGSSVGLLALLITLALITLALYKLGFFKQYKDMSEGGPPGAEPQ 1153

RESULT 9  
US-09-688-307A-3  
; Sequence 3, Application US/09688307A  
; Patent No. 6432404  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: NO. 6432404el Human Beta-2  
; FILE REFERENCE: 27866/36646  
; CURRENT APPLICATION NUMBER: US/09/688.307A  
; CURRENT FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: 09/193,043  
; PRIOR FILING DATE: 1998-11-16  
; PRIOR APPLICATION NUMBER: 08/605,672  
; PRIOR FILING DATE: 1996-02-22  
; PRIOR APPLICATION NUMBER: 08/173,497  
; PRIOR FILING DATE: 1993-12-23  
; PRIOR APPLICATION NUMBER: 08/286,889  
; PRIOR FILING DATE: 1994-08-05  
; PRIOR APPLICATION NUMBER: 08/362,652  
; PRIOR FILING DATE: 1994-12-21  
; PRIOR APPLICATION NUMBER: 08/943,363  
; PRIOR FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-688-307A-3

Query Match 99.8%; Score 5862; DB 4; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANQSGLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANQSGLYQCDYSTGSCPEI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQOPQK 136  
Qy 121 FPALRGCPQEDSDIAFLVDGSGSIIPHDPRMKEFVSTVMEQKKSKTLFSLMOYSEEF 180  
Db 137 FPALRGCPQEDSDIAFLVDGSGSIIPHDPRMKEFVSTVMEQKKSKTLFSLMOYSEEF 196  
Qy 181 RIHFTFKFQNNPNRSLIKPIQTLLGRTHATGIRKVVRELFTNGASKNAFKILILI 240  
Db 197 RIHFTFKFQNNPNRSLIKPIQTLLGRTHATGIRKVVRELFTNGASKNAFKILILI 256  
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKROELNTVASKPRDHFVQIN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKROELNTVASKPRDHFVQIN 316  
Qy 301 NFEALKTIONOLREKIFAIETGQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 360

; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-193-043-3

Query Match 99.8%; Score 5862; DB 3; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANQSGLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVAANQSGLYQCDYSTGSCPEI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLPGSNLRQOPQK 136  
Qy 121 FPALRGCPQEDSDIAFLVDGSGSIIPHDPRMKEFVSTVMEQKKSKTLFSLMOYSEEF 180  
Db 137 FPALRGCPQEDSDIAFLVDGSGSIIPHDPRMKEFVSTVMEQKKSKTLFSLMOYSEEF 196  
Qy 181 RIHFTFKFQNNPNRSLIKPIQTLLGRTHATGIRKVVRELFTNGASKNAFKILILI 240  
Db 197 RIHFTFKFQNNPNRSLIKPIQTLLGRTHATGIRKVVRELFTNGASKNAFKILILI 256  
Qy 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKROELNTVASKPRDHFVQIN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKROELNTVASKPRDHFVQIN 316  
Qy 301 NFEALKTIONOLREKIFAIETGQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 360  
Db 317 NFEALKTIONOLREKIFAIETGQTGSSSSFEHMSQEGFSAAITSNGLPSTVGSYDWAG 376  
Qy 361 GVFLYTSKSTFFINNRVDSMDNAYLGAAAILRNVRQSLVILGAPYQHIGLVAMER 420  
Db 377 GVFLYTSKSTFFINNRVDSMDNAYLGAAAILRNVRQSLVILGAPYQHIGLVAMER 436  
Qy 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIIGAPHYEQTRGGQSVCP 480  
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIIGAPHYEQTRGGQSVCP 496  
Qy 481 PRGORARWQCDVLYGEQGPWGRFGAALTVDLVNGDKLTDVAIGAPBEDNRGAVYLF 540  
Db 497 PRGORARWQCDVLYGEQGPWGRFGAALTVDLVNGDKLTDVAIGAPBEDNRGAVYLF 556  
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGQGHVLLRSQ 616  
Qy 601 PVLVRKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 660  
Db 617 PVLVRKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLREGQIQSVVT 676  
Qy 661 YDLALDSGRPHSAVFNFTKNSTRQTQVLGLTQTCETLKLQPNCTIEDPVPVILRLNF 720  
Db 677 YDLALDSGRPHSAVFNFTKNSTRQTQVLGLTQTCETLKLQPNCTIEDPVPVILRLNF 736  
Qy 721 SLVGTPLSAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAGNLRPVLAEDAQRLLFTALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
Db 797 GPREFNVTVRNDGDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
Qy 841 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTFTNITFDVDSKASLGNKLLKANVTSENNMPRTNTEF 916  
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVMOHQVSNLGRSLPISLVFLVPV 960

Db 317 NFEALKTIONOLREKIFAIEGTQTGSSSSFFHEMSQEGFSAATSNGLSTVGSYDWAG 376  
Qy 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436  
Qy 421 QNTGWSNANVKTQTIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIEQTRGGQVSVCP 480  
Db 437 QNTGWSNANVKTQTIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIEQTRGGQVSVCP 496  
Qy 481 PRGORARWQCDVLYGQOGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540  
Db 497 PRGORARWQCDVLYGQOGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 556  
Qy 541 HGTSGSGISPSHSORIASGSLSPRLQYFGQSLGGQDLTMDGLVLTGAGQHVLLRSQ 600  
Db 557 HGTSGSGISPSHSORIASGSLSPRLQYFGQSLGGQDLTMDGLVLTGAGQHVLLRSQ 616  
Qy 601 PVLVRKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660  
Db 617 PVLVRKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676  
Qy 661 YDLALDSGRPHSAVRNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720  
Db 677 YDLALDSGRPHSAVRNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 736  
Qy 721 SLVGTPLSAFNLPRVLAEDAQRFLFTALFFPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLPRVLAEDAQRFLFTALFFPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 856  
Qy 841 SGALKSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 900  
Db 857 SGALKSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 916  
Qy 901 QLELPVKAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLSLVFLVPV 960  
Db 917 QLELPVKAVYVMTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSPLSLVFLVPV 976  
Qy 961 RLNQTVIWDROPVTFSENLSSTCHTKERLPHSDPLAELRKAAPVNCVSIQVQRIQCIP 1020  
Db 977 RLNQTVIWDROPVTFSENLSSTCHTKERLPHSDPLAELRKAAPVNCVSIQVQRIQCIP 1036  
Qy 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQOGAFVRSQTET 1080  
Db 1037 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPQOGAFVRSQTET 1096  
Qy 1081 KVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKROYKDMMSBGGPPGABPQ 1137  
Db 1097 KVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKROYKDMMSBGGPPGABPQ 1153

## RESULT 10

US-09-350-259-3  
; Sequence 3, Application US/09350259  
; Patent No. 6620915  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6620915el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; EARLIER FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-3

Query Match 99.8%; Score 5862; DB 4; Length 1153;  
Best Local Similarity 99.3%; Pred. No. 0;  
Matches 1129; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVCAPOEIVAAORGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLOGSRVVVCAPOEIVAAORGSLYQCDYSTGSCPEI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKLCFLFGSNLRQPOK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVVKLCFLFGSNLRQPOK 136  
Qy 121 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEPISTWMEQLKKSKTLFSLMOYSEBF 180  
Db 137 FPEALRGCPQEDSDIAFLVDGSGSIIPHDPRRAKEPISTWMEQLKKSKTLFSLMOYSEBF 196  
Qy 181 RHIFTFKEFQNNPNRSLIKPITQLGRTHATGIRKVVRELFINITNGARKNAFKLILVI 240  
Db 197 RHIFTFKEFQNNPNRSLIKPITQLGRTHATGIRKVVRELFINITNGARKNAFKLILVI 256  
Qy 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFRSEKSRQBLNVTASKPRDHVFOIN 300  
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGDAFRSEKSRQBLNVTASKPRDHVFOIN 316  
Qy 301 NFEALKTIONOLREKIFAIEGTQTGSSSSFFHEMSQEGFSAATSNGLSTVGSYDWAG 360  
Db 317 NFEALKTIONOLREKIFAIEGTQTGSSSSFFHEMSQEGFSAATSNGLSTVGSYDWAG 376  
Qy 361 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFFINMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYOHIGLVAMFR 436  
Qy 421 QNTGWSNANVKTQTIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIEQTRGGQVSVCP 480  
Db 437 QNTGWSNANVKTQTIGAYFGASLCSVDVDSNGSTDLVLIGAPHYIEQTRGGQVSVCP 496  
Qy 481 PRGORARWQCDVLYGQOGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 540  
Db 497 PRGORARWQCDVLYGQOGPWGFRGAALTIVLGVNGDKLTDVAIGAPGEDNRGAVYLF 556  
Qy 541 HGTSGSGISPSHSORIASGSLSPRLQYFGQSLGGQDLTMDGLVLTGAGQHVLLRSQ 600  
Db 557 HGTSGSGISPSHSORIASGSLSPRLQYFGQSLGGQDLTMDGLVLTGAGQHVLLRSQ 616  
Qy 601 PVLVRKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660  
Db 617 PVLVRKAIEMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676  
Qy 661 YDLALDSGRPHSAVRNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720  
Db 677 YDLALDSGRPHSAVRNETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 736  
Qy 721 SLVGTPLSAFNLPRVLAEDAQRFLFTALFFPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFNLPRVLAEDAQRFLFTALFFPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840  
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 856  
Qy 841 SGALKSTCSINHPFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPRTNKTEF 900

857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 916  
901 QLELPVKYAVYVMVTSYKYNFTASENTRVMOHQYQVSNLQKSLPISLVFLVPV 960  
917 QLELPVKYAVYVMVTSYKYNFTASENTRVMOHQYQVSNLQKSLPISLVFLVPV 976  
961 RLNQTIVDRPQVTFSENLSTCHTKERLPKSHDPLAELRKAPVNVNCSIAVCQIQCDIP 1020  
977 RLNQTIVDRPQVTFSENLSTCHTKERLPKSHDPLAELRKAPVNVNCSIAVCQIQCDIP 1036  
1021 PFGIOEENATLKGNLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1080  
1037 PFGIOEENATLKGNLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1096  
1081 KVEPEVENPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPO 1137  
1097 KVEPEVENPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPO 1153

RESULT 11  
US-08-476-062A-43  
; Sequence 43, Application US/08476062A  
; Patent No. 5877275  
; GENERAL INFORMATION:  
; APPLICANT: Arnaout, M. Amin  
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY  
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: US  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/476,062A  
; FILING DATE: 07-JUN-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/216,081  
; FILING DATE: 21-MAR-1994  
; APPLICATION NUMBER: 07/637,830  
; FILING DATE: 04-JAN-1991  
; APPLICATION NUMBER: 07/539,842  
; FILING DATE: 18-JUN-1990  
; APPLICATION NUMBER: 07/212,573  
; FILING DATE: 28-JUN-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00786/068003  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELE: 200154  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1152 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal  
; US-08-476-062A-43

Query Match 99.2%; Score 5831.5; DB 2; Length 1152;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

1 FNLDTENAMTFQENARFGQSVVOLQGSRRVVVGAPOEIVAAQNRGSILYQCDYSTGSCPEI 60  
17 FNLDTENAMTFQENARFGQSVVOLQGSRRVVVGAPOEIVAAQNRGSILYQCDYSTGSCPEI 76  
61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQTCSENTYVVKGLCFGLFNSLRQOPQK 120  
77 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHTQTCSENTYVVKGLCFGLFNSLRQOPQK 136  
121 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDFFRAKEFIISTVMEOQKKSTLSLMQYSBEF 180  
137 FPEALRGCPQSDSDIAFLVDGSGSIIIPHDFFRAKEFIISTVMEOQKKSTLSLMQYSBEF 196  
181 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILLI 240  
197 RIHFTFKFQNNPNRSLIKPITQLLGRTHATGIRKVVRELFINITNGARKNAFKILLI 256  
241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELMTVASKPRDHVFN 300  
257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELMTVASKPRDHVFN 316  
301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLLSTVGSYDWAG 360  
317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAITNSGPLLLSTVGSYDWAG 376  
361 GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAAIIIRNRVQSLVIGAPYQHIGLVAMFR 420  
377 GVFLYTSKEKSTFINMTRVDSMDNAYLGYAAAIIIRNRVQSLVIGAPYQHIGLVAMFR 436  
421 QNTGMWESNANVKTQIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYVEQTRGQSVQVCL 480  
437 QNTGMWESNANVKTQIGAYFGASLCSDVDVDSNGSTDLVLIGAPHYVEQTRGQSVQVCL 496  
481 PRGORARWOCDAVLYGEOQPGWGRGAALTIVLGDVNGDKLTDVAIGAPBEDNRGAYLF 540  
497 PRG-RARWOCDAVLYGEOQPGWGRGAALTIVLGDVNGDKLTDVAIGAPBEDNRGAYLF 555  
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 615  
601 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGIQSVVT 660  
616 PVLVRKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDLREGIQSVVT 675  
661 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSVTLRLNF 720  
676 YDLALDSGRPHSRAVFNETKNSTRQTQVLGLTQTCETLKLQLPNCIEDVPSVTLRLNF 735  
721 SLVGTPLSAFNGLRPVLAEDAQRFTALPFPEKNCNDNICODDLSITPFSMLDCLVVG 780  
736 SLVGTPLSAFNGLRPVLAEDAQRFTALPFPEKNCNDNICODDLSITPFSMLDCLVVG 795  
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 840  
796 GPRESNVTVVRNDGEDSYRTQVTFPPDLDSYRKVSTLQNRQSRWRLACESASSTEV 855  
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900  
856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 915  
901 QLELPVKYAVYVMVTSYKYNFTASENTRVMOHQYQVSNLQKSLPISLVFLVPV 960  
916 QLELPVKYAVYVMVTSYKYNFTASENTRVMOHQYQVSNLQKSLPISLVFLVPV 975  
961 RLNQTIVDRPQVTFSENLSTCHTKERLPKSHDPLAELRKAPVNVNCSIAVCQIQCDIP 1020  
976 RLNQTIVDRPQVTFSENLSTCHTKERLPKSHDPLAELRKAPVNVNCSIAVCQIQCDIP 1035  
1021 PFGIOEENATLKGNLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1080  
1036 PFGIOEENATLKGNLSPDWYIKTSHNHLIVSTAEILFNDSVFTLLPQOGAFVRSQDET 1095  
1081 KVEPEVENPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGABPO 1137

Db 1096 KVEPFVFNPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1152

RESULT 12  
PCT-US96-01314-43  
Sequence 43, Application PC/TUS9601314  
GENERAL INFORMATION:  
APPLICANT: M. Amin Arnaut  
TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01314  
FILING DATE: 30-JAN-96  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/380,167  
FILING DATE: 30-JAN-95  
ATTORNEY/AGENT INFORMATION:  
NAME: John W. Freeman  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/267001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1152  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
PCT-US96-01314-43

Query Match 99.2%; Score 5831.5; DB 5; Length 1152;  
Best Local Similarity 99.0%; Pred. No. 0;  
Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARGFGOSVVOLOGSRVVGAPQIVAAANQSGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGOSVVOLOGSRVVGAPQIVAAANQSGSLYQCDYSTGSCPEI 76  
QY 61 RLOQVPEAVNMSLGLSLAATSPQLLACGPTVHTQTCSENTYVKGCLFGLSGNLROQPOK 120  
Db 77 RLOQVPEAVNMSLGLSLAATSPQLLACGPTVHTQTCSENTYVKGCLFGLSGNLROQPOK 136  
QY 121 FPALRGCPQEDSDIAPLVDSGSIIPHDFRAKEPISTVMEQIKKSKTIFSLMQYSEEF 180  
Db 137 FPALRGCPQEDSDIAPLVDSGSIIPHDFRAMEKEFVSTVMEQIKKSKTIFSLMQYSEEF 196  
QY 181 RIHFTKPFQNNPRSLKPIITQLLGRTHATGIRKVVRELNTINGARKNAFKILILI 240  
Db 197 RIHFTKPFQNNPRSLKPIITQLLGRTHATGIRKVVRELNTINGARKNAFKILIVI 256  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRBELNTVASKPPRDHVFOIN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFSEKSRBELNTVASKPPRDHVFOIN 316  
QY 301 NFEALKTIONOLREKIFAIEGTGTGSSSFHEMSQEGFSAITSNGLPLSTVGSYDWAG 360  
Db 317 NFEALKTIONOLREKIFAIEGTGTGSSSFHEMSQEGFSAITSNGLPLSTVGSYDWAG 376

QY 361 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRVOHIGLVAMFR 420  
Db 377 GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNRVQSLVGLAPRVOHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQSVCP 480  
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEYQTRGGQSVCP 496  
QY 481 PRGORARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDAIAGPGEEDNRGAYLYF 540  
Db 497 PRG-RARWQCDVLYGEOGQPMGRFGAALTVLGDVNGDKLTDAIAGPGEEDNRGAYLYF 555  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
Db 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTVGAQGHVLLRSQ 615  
QY 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 616 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 675  
QY 661 YDLALDSGRPHSRVFNENETKSTRQTVGLGTQTCETLKLQIPNCIEDPVPSPVLRNLF 720  
Db 676 YDLALDSGRPHSRVFNENETKSTRQTVGLGTQTCETLKLQIPNCIEDPVPSPVLRNLF 735  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFPEKNCNDNI CODDLSITFSMSLDCLVVG 780  
Db 736 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPFPEKNCNDNI CODDLSITFSMSLDCLVVG 795  
QY 781 GPREENVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASSTEV 840  
Db 796 GPREENVTVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRWELACESASSTEV 855  
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENMMPTNKTEF 900  
Db 856 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGKLLKANVTSENMMPTNKTEF 915  
QY 901 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMOHOYOVSNLQSRSLPSLVFLVPV 960  
Db 916 QLELPVKYAVVMVVTSHGVSTKYLNFNTASENTSRVMOHOYOVSNLQSRSLPSLVFLVPV 975  
QY 961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1020  
Db 976 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP 1035  
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAETILFNDVSFTLLPGQAFVRSQTE 1080  
Db 1036 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAETILFNDVSFTLLPGQAFVRSQTE 1095  
QY 1081 KVEPFVFNPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1137  
Db 1096 KVEPFVFNPLPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGGPPGAEPQ 1152

RESULT 13  
5424399-2  
Patent No. 5424399  
APPLICANT: ARNAOUT, M. AMIN  
TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/78,871  
FILING DATE: 16-JUN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 539,842  
FILING DATE: 18-JUN-1990  
APPLICATION NUMBER: 212,573  
FILING DATE: 28-JUN-1988  
SEQ ID NO: 2;  
LENGTH: 1152  
5424399-2

Query Match 99.2%; Score 5831.5; DB 6; Length 1152;

		Best Local Similarity 99.0%; Pred. No. 0;		Matches 1126; Conservative 7; Mismatches 3; Indels 1; Gaps 1;	
Qy	1	FNLDTENAMTFOENARFGQSVVQLOQSRVVVGAPOEIVANQSGSLYQCDYSTGSCPEI	60		
Db	17	FNLDTENAMTFOENARFGQSVVQLOQSRVVVGAPOEIVANQSGSLYQCDYSTGSCPEI	76		
Qy	61	RLQVPVAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQDQOK	120		
Db	77	RLQVPVAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQDQOK	136		
Qy	121	FPEALRGCPQSDSDIAFLVDGSGSLIIPHDPPRAKEFTSTVMEQKKSKTLFSLMOYSEEF	180		
Db	137	FPEALRGCPQSDSDIAFLVDGSGSLIIPHDPPRAKEFTSTVMEQKKSKTLFSLMOYSEEF	196		
Qy	181	RIHFTFKFQNNPNRSLIKPTITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI	240		
Db	197	RIHFTFKFQNNPNRSLIKPTITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI	256		
Qy	241	TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSKROELNTVASKPPRDHVFQIN	300		
Db	257	TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAPRSEKSKROELNTVASKPPRDHVFQIN	316		
Qy	301	NFEALKTTQNLREKIFAIEGTQTSSEFEHMSQEGFSAAITNSGPLLSTVGSYDQWAG	360		
Db	317	NFEALKTTQNLREKIFAIEGTQTSSEFEHMSQEGFSAAITNSGPLLSTVGSYDQWAG	376		
Qy	361	GVFLYTSKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVGLGAPYQHIGLVAMPR	420		
Db	377	GVFLYTSKSTFINNTRVDSMDNDAYLGAAAILRNVRQSLVGLGAPYQHIGLVAMPR	436		
Qy	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP	480		
Db	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGQSVCP	496		
Qy	481	PRGORARQCDVLYGEQGPWGRFCAALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLF	540		
Db	497	PRG-RARWQCDVLYGEQGPWGRFCAALTVLGVNGDKLTDVAIGAPGEEDNRGAVYLF	555		
Qy	541	HGTSGSGISPSHSORLAGSKSLPRLOVFGQSLGGQDLTMDGLVDLTVGAGHVLRLSQ	600		
Db	556	HGTSGSGISPSHSORLAGSKSLPRLOVFGQSLGGQDLTMDGLVDLTVGAGHVLRLSQ	615		
Qy	601	PVLVRKALMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT	660		
Db	616	PVLVRKALMEFNPREVARNVFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT	675		
Qy	661	YDLALDSGRPHSAVFNETKNSTRQOVLTGTLQTCETLKLQPNCLIEDPVSPVLRLNF	720		
Db	676	YDLALDSGRPHSAVFNETKNSTRQOVLTGTLQTCETLKLQPNCLIEDPVSPVLRLNF	735		
Qy	721	SLVGTPLSAFGLNRPVLAEDAQRULTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG	780		
Db	736	SLVGTPLSAFGLNRPVLAEDAQRULTALFPPEKNCNDNICQDDLSITFSFMSLDCLVVG	795		
Qy	781	GPREFNVTVVRNDGDSRYRTQVTFPPPLDLSYKXVSTLQNRQSRQSRWLACESASSTEV	840		
Db	796	GPRESNVTVVRNDGDSRYRTQVTFPPPLDLSYKXVSTLQNRQSRQSRWLACESASSTEV	855		
Qy	841	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPNRNTKTEF	900		
Db	856	SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENNPNRNTKTEF	915		
Qy	901	QLELPVKYAVVMVTSVSHGVSTKYLNFNTASNTSRVMQHOYQVSNLQGRSLPISLVFLVPV	960		
Db	916	QLELPVKYAVVMVTSVSHGVSTKYLNFNTASNTSRVMQHOYQVSNLQGRSLPISLVFLVPV	975		
Qy	961	RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP	1020		
Db	976	RLNQTVIWRDPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP	1035		
Qy	1021	FFGIQIEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSVFTLLPQGAFAVRSQTEI	1080		
Db	1036	FFGIQIEEFNATLKGNSLSDWYIKTSHNHLIVSTAEILFNDSVFTLLPQGAFAVRSQTEI	1095		
Qy	1081	KVPEFVEPNPLPLTVGSSVGLLILALITAAALYKLGFFKQYKDMMEGGPPGAEPQ	1137		
Db	1096	KVPEFVEPNPLPLTVGSSVGLLILALITAAALYKLGFFKQYKDMMEGGPPGAEPQ	1152		
RESULT 14					
US-08-476-062A-44					
; Sequence 44: Application US/08476062A					
; Patent No. 5877275					
; GENERAL INFORMATION:					
; APPLICANT: Arnaout, M. Amin					
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY					
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS					
; NUMBER OF SEQUENCES: 53					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Fish & Richardson P.C.					
; STREET: 225 Franklin Street					
; CITY: Boston					
; STATE: MA					
; COUNTRY: US					
; ZIP: 02110-2804					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Diskette					
; COMPUTER: IBM Compatible					
; OPERATING SYSTEM: Windows95					
; SOFTWARE: Fast-SEQ for Windows Version 2.0					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/476,062A					
; FILING DATE: 07-JUN-1995					
; PRIOR APPLICATION DATA:					
; APPLICATION NUMBER: 08/216,081					
; FILING DATE: 21-MAR-1994					
; APPLICATION NUMBER: 07/637,830					
; FILING DATE: 04-JAN-1991					
; APPLICATION NUMBER: 07/539,842					
; FILING DATE: 18-JUN-1990					
; APPLICATION NUMBER: 07/212,573					
; FILING DATE: 28-JUN-1988					
; ATTORNEY/AGENT INFORMATION:					
; NAME: Freeman, John W.					
; REGISTRATION NUMBER: 29,066					
; REFERENCE/DOCKET NUMBER: 00786/068003					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: 617/542-5070					
; TELEFAX: 617/542-8906					
; TELEX: 200154					
; INFORMATION FOR SEQ ID NO: 44:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 1163 amino acids					
; TYPE: amino acid					
; TOPOLOGY: linear					
; MOLECULAR TYPE: protein					
US-08-476-062A-44					
Query Match 59.0%; Score 3469; DB 2; Length 1163;					
Best Local Similarity 61.0%; Pred. No. 7e-284;					
Matches 689; Conservative 141; Mismatches 293; Indels 6; Gaps 4;					
Qy	1	FNLDTENAMTFOENARFGQSVVQLOQSRVVVGAPOEIVANQSGSLYQCDYSTGSCPEI	60		
Db	20	FNLDTENAMTFOENARFGQSVVQLOQSRVVVGAPOEIVANQSGSLYQCDYSTGSCPEI	79		
Qy	61	RLQVPVAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQDQOK	120		
Db	80	GLQVPPEAVNMVSLGLSLAATTSPPQLLACGPTVHQTCSNTYVYKGLCFLFGSNLRQDQOK	137		
Qy	121	FPEALRGCPQSDSDIAFLVDGSGSLIIPHDPPRAKEFTSTVMEQKKSKTLFSLMOYSEEF	180		
Db	138	LPVSRQECPRQEQDIFVLIDGSGSSRNFNATMMNFVRAVISQFORPSTQFSLMOFSNKF	197		
Qy	181	RIHFTFKFQNNPNRSLIKPTITQLLGRTHATGIRKVVRELFNITNGARKNAFKILILI	240		





Tue Nov 9 12:56:09 2004

QY	541	HGTS	GGIS	PSHS	QRI	AGSK	LS	PL	QY	FG	QS	LS	GG	DL	T	M	D	G	L	V	D	L	T	V	G	A	Q	H	V	L	L	R	S	Q	600																											
Db	557	HG	V	L	G	P	S	I	S	P	S	H	S	Q	R	I	A	G	S	S	R	L	Q	Y	F	G	Q	L	T	D	G	L	V	D	L	A	V	G	A	Q	H	V	L	L	R	T	R	616														
QY	601	P	V	L	R	V	K	A	I	M	E	N	F	P	E	V	A	R	N	F	E	C	N	D	O	V	V	K	E	A	G	E	V	R	V	C	L	H	V	Q	K	S	T	R	D	R	E	G	O	I	Q	S	V	T	660							
Db	617	P	V	L	R	V	K	A	I	M	E	N	F	P	E	V	A	R	N	F	E	C	N	D	O	V	V	K	E	A	G	E	V	R	V	C	L	H	V	Q	K	S	T	R	D	R	E	G	O	I	Q	S	V	T	676							
QY	661	Y	D	L	A	D	S	G	R	P	H	S	R	A	V	F	N	E	T	K	N	S	T	R	Q	T	O	V	L	G	T	O	T	C	E	T	K	L	Q	L	P	N	C	I	E	D	P	V	S	P	I	V	I	R	N	F	720					
Db	677	L	D	L	A	D	P	G	R	L	S	P	R	A	T	F	Q	E	T	K	N	S	T	R	Q	T	O	V	L	G	T	O	T	C	E	T	K	L	Q	L	P	N	C	I	E	D	P	V	S	P	I	V	I	R	N	F	736					
QY	721	S	I	V	G	T	P	L	S	A	F	G	N	L	R	P	V	L	A	E	D	A	O	R	L	F	T	A	L	P	P	F	E	K	N	C	G	N	D	N	I	C	O	D	D	L	S	I	T	F	S	E	M	S	L	D	C	L	V	V	G	780
Db	737	T	I	V	G	K	P	L	A	F	R	N	L	R	P	M	L	A	L	A	Q	R	Y	E	T	A	S	L	P	F	E	K	N	C	G	A	D	H	I	C	O	D	N	L	G	I	S	F	S	F	F	G	L	K	S	L	L	V	G	796		
QY	781	G	P	R	E	N	V	T	V	T	R	N	D	G	E	D	S	V	R	T	O	V	T	F	P	P	L	D	L	S	V	R	K	V	S	T	L	Q	N	O	R	S	O	R	S	R	L	A	C	E	S	A	S	T	E	V	840					
Db	797	S	N	L	E	N	A	E	V	M	W	N	D	G	E	D	S	G	T	T	I	T	S	H	P	A	G	L	S	V	R	Y	V	A	E	G	O	K	O	G	L	R	S	L	H	L	T	C	D	S	A	P	V	G	--	854						
QY	841	S	G	A	L	K	S	T	C	S	I	N	H	P	I	P	E	N	S	E	V	T	F	N	I	T	F	D	V	S	K	A	S	L	G	N	K	L	L	K	A	N	V	T	S	E	N	N	M	P	R	I	N	K	T	E	F	900				
Db	855	S	Q	T	W	S	T	C	R	I	N	H	L	I	F	R	G	A	I	T	F	L	A	T	F	D	V	S	P	K	A	V	L	G	R	L	L	L	T	A	N	V	S	S	E	N	T	P	R	I	S	K	T	T	F	914						
QY	901	Q	L	E	L	P	V	K	A	V	Y	T	V	S	S	H	E	O	F	T	K	Y	L	N	F	S	E	S	E	K	E	H	V	A	M	H	R	Y	Q	V	N	N	I	G	O	R	D	L	P	V	S	I	N	F	W	P	959					
Db	915	Q	L	E	L	P	V	K	A	V	Y	T	V	S	S	H	E	O	F	T	K	Y	L	N	F	S	E	S	E	K	E	H	V	A	M	H	R	Y	Q	V	N	N	I	G	O	R	D	L	P	V	S	I	N	F	W	P	974					
QY	960	V	R	L	N	O	T	V	I	W	D	R	P	O	V	T	F	S	E	N	L	S	T	C	H	T	K	E	R	L	P	S	H	S	D	F	L	A	E	L	R	K	A	P	V	V	N	C	S	I	A	V	C	O	R	I	O	C	D	I	1019	
Db	975	V	E	L	N	O	E	A	V	M	D	V	E	S	H	P	O	N	F	S	L	R	C	S	E	K	I	A	P	P	A	S	D	F	L	A	H	I	Q	K	N	P	V	L	D	C	S	I	A	G	L	R	F	R	C	D	V	1034				
QY	1020	P	F	T	G	I	O	E	E	F	N	A	T	L	K	G	N	L	S	F	D	W	Y	I	K	T	S	H	N	H	L	L	I	V	S	T	A	E	I	L	P	N	D	S	V	F	T	L	L	P	O	G	A	F	V	R	S	O	T	E	1079	
Db	1035	P	S	E	S	V	O	E	E	L	D	F	T	L	K	G	N	L	S	F	G	W	R	Q	I	L	Q	K	V	S	V	S	V	A	E	I	T	E	D	T	S	V	I	S	O	L	P	O	G	E	A	F	M	R	A	Q	T	1094				
QY	1080	T	K	V	E	P	F	V	P	N	P	L	I	V	G	S	S	V	G	L	L	L	L	L	I	T	A	L	I	T	A	A	L	K	G	L	F	F	K	R	O	Y	K	D	M	M	S	E	1128													
Db	1095	T	V	L	E	K	Y	K	V	H	N	P	T	L	I	V	G	S	S	I	G	G	L	L	L	L	I	T	A	V	L	Y	K	V	G	F	F	R	Q	Y	K	E	M	M	E	E	1143															

Search completed: November 9, 2004, 12:21:09  
Job time : 28 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:37 ; Search time 27.25 Seconds  
(without alignments)  
4014.622 Million cell updates/sec

Title: US-09-902-481b-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPCAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:\*

1: Pir1:\*

2: Pir2:\*

3: Pir3:\*

4: Pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5868	99.9	1153	1 RWHU1B	cell surface glyco
2	4476	76.2	1153	2 S00551	leukocyte surface
3	3483	59.3	1163	1 RWHU1C	cell surface glyco
4	1548.5	26.4	1170	2 S03308	cell surface glyco
5	1533.5	26.1	1163	2 I56126	lymphocyte fuction
6	1149	19.6	1179	2 A53213	integrin alpha-E c
7	1102.5	18.8	1151	2 A43226	integrin alpha-1 c
8	1084	18.5	1170	2 I45914	integrin alpha-1 c
9	1072	18.2	1178	2 S44142	integrin alpha 2 s
10	1069	18.2	1181	2 A33998	integrin alpha-2 c
11	1060	18.0	1180	2 A35854	integrin alpha-1 c
12	663	11.3	1039	2 A41131	lymphocyte-Peyer's
13	633	10.8	1038	2 S06046	integrin alpha-4 c
14	630	10.7	1035	2 I58409	integrin alpha-4 c
15	614.5	10.5	1041	2 T31437	integrin alpha-9 c
16	579.5	9.9	1054	2 JC7294	integrin alpha cha
17	572.5	9.7	1051	2 A35761	cell surface glyco
18	567.5	9.7	1053	2 I55534	integrin alpha-6 c
19	555.5	9.5	1053	2 S44250	integrin alpha-6 c
20	543.5	9.3	1034	2 A36108	integrin alpha-5 c
21	535	9.1	1044	2 T10050	integrin alpha-5 c
22	532	9.1	1049	2 A27079	fibronectin recept
23	532	9.1	1073	2 B36429	integrin alpha-6 c
24	530.5	9.0	1072	2 A38457	integrin alpha-6 c
25	529.5	9.0	1051	2 A40021	integrin alpha-3 c
26	526	9.0	1048	2 A27421	integrin alpha-5 c
27	525.5	8.9	1091	2 A41543	integrin alpha-6 c
28	517	8.8	1044	2 S16516	integrin alpha-8 c
29	496.5	8.5	1146	2 S40311	integrin - fruit f

#### ALIGNMENTS

##### RESULT 1

###### RWHU1B

cell surface glycoprotein CD11b precursor [validated] - human

N;Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac-1  
eukocyte integrin alpha chain; neutrophil adherence receptor alphaM chain  
C;Species: Homo sapiens (man)

C;Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 09-Jul-2004

C;Accession: A31108; A28915; A41600; A30892; A32218; A46526; A36091; I52567

R;Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A.

J. Biol. Chem. 263, 12403-12411, 1988

A;Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b).

A;Reference number: A31108; MUID:88315033; PMID:2457584

A;Accession: A31108

A;Molecule type: mRNA

A;Residues: 1-1153 <COR>

A;Cross-references: UNIPROT:P11215; GB:J03925; NID:gl87284; PIDN:AAAS5944.1; PID:g307148

A;Note: part of this sequence was confirmed by protein sequencing

R;Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G.

J. Cell Biol. 106, 2153-2158, 1988

A;Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor N

A;Reference number: A28915; MUID:88257215; PMID:2454931

A;Accession: A28915

A;Molecule type: mRNA

A;Residues: 1-499,501-965,'P',967-1153 <ARN>

A;Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:gl86935; PIDN:AAAS5944.1

A;Note: the authors translated the codon TAC for residue 1129 as Thr

A;Note: part of this sequence, including the amino end of the mature protein, was confirmed

R;Shelley, C.S.; Arnaout, M.A.

Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A;Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regulated

A;Reference number: A41600; MUID:92073318; PMID:1683702

A;Accession: A41600

A;Molecule type: DNA

A;Residues: 1-9 <SH>

A;Cross-references: GB:M76724; NID:gl80018; PIDN:AAAS8410.1; PID:g553215

R;Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G.

Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A;Title: Molecular cloning of the alpha-subunit of human and guinea pig leukocyte adhesion

A;Reference number: A94193; MUID:88190151; PMID:2833753

A;Accession: A30892

A;Molecule type: mRNA

A;Residues: 917-1042 <AR2>

A;Cross-references: GB:M18044

R;Hickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J.

Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A;Title: cDNA sequence for the alphaM subunit of the human neutrophil adherence receptor

A;Reference number: A32218; MUID:89098893; PMID:2563162

A;Accession: A32218

A;Molecule type: mRNA

A;Residues: 9-1153 <HIC>

A;Cross-references: GB:J04145; NID:gl89068; PIDN:AAAS59903.1; PID:g386975

A;Note: part of this sequence was confirmed by protein sequencing  
 R;Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G.  
 J. Immunol. 150, 480-490, 1993  
 A;Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-in  
 n during evolution.  
 A;Reference number: A46526; MUID:93123748; PMID:8419480  
 A;Accession: A46526  
 A;Status: not compared with conceptual translation  
 A;Molecule type: DNA  
 A;Residues: 1-499,501-1153 <FILE>  
 A;Cross-references: GB:552227; NID:g263047; PIDN:AA24821.1; PID:g263049  
 A;Note: the last three bases of intron 13, CAG, are included in some but not all mature  
 A;Note: sequence extracted from NCBI backbone (NCBIP:121963)  
 R;Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A.  
 Biochim. Biophys. Acta 874, 368-371, 1986  
 A;Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across sp  
 A;Reference number: A90664; MUID:87076671; PMID:3539202  
 A;Accession: A26091  
 A;Molecule type: protein  
 A;Residues: 17-31 <PIE>  
 A;Experimental source: granulocytes  
 R;Pahl, H.L.; Roemarin, A.G.; Tenen, D.G.  
 Blood 79, 865-870, 1992  
 A;Title: Characterization of the myeloid-specific CD11b promoter.  
 A;Reference number: 152567; MUID:92144986; PMID:1346576  
 A;Accession: 152567  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-9 <RES>  
 A;Cross-references: GB:M84477; NID:g180184; PIDN:AA51960.1; PID:g553219  
 C;Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1  
 C;Genetics:  
 A;Gene: GDB:ITGAM; CR3A  
 A;Cross-references: GDB:120599; OMIM:120980  
 A;Map position: 16p11.2-16p11.2  
 A;Note: promoter contains a GATA motif and two Sp1 consensus binding sites  
 C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
 C;Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag  
 F;1-16/Domain: signal sequence #status predicted <SIG>  
 F;17-1153/Product: cell surface glycoprotein CD11b #status experimental <MAT>  
 F;17-1108/Domain: extracellular #status predicted <EXT>  
 F;148-318/Domain: von Willebrand factor type A repeat homology <VWA2>  
 F;465-473/Region: calcium/magnesium binding #status predicted  
 F;530-538/Region: calcium/magnesium binding #status predicted  
 F;593-601/Region: calcium/magnesium binding #status predicted  
 F;1109-1134/Domain: transmembrane #status predicted <TM>  
 F;1135-1153/Domain: intracellular #status predicted <INT>  
 F;86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 99.9%; Score 5868; DB 1; Length 1153;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANQSGSLYQCDYSTGCEPI 60  
 Db 17 FNLDENAMTFQENARGFGQSVVQLQSGRVVVGAPQEIIVANQSGSLYQCDYSTGCEPI 76  
 Qy 61 RLQVPVAVNMSGLSLAATSPQLLACGTVHTQTSNTYVKGCLFCLFGSNLRQDPQK 120  
 Db 77 RLQVPVAVNMSGLSLAATSPQLLACGTVHTQTSNTYVKGCLFCLFGSNLRQDPQK 136  
 Qy 121 FPEALRCPCQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 180  
 Db 137 FPEALRCPCQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196  
 Qy 181 RIHFTTFEQNNPRSLVKEITQLGRTHATGVRKVIKRELLNITGARKNAFKILIVI 240  
 Db 197 RIHFTTFEQNNPRSLVKEITQLGRTHATGVRKVIKRELLNITGARKNAFKILIVI 256  
 Qy 241 TDGKFGDPLGVEDVIPADREGVIRVYGVGDAPFSEKSKQELNITASKPRDHVQVN 300  
 Db 257 TDGKFGDPLGVEDVIPADREGVIRVYGVGDAPFSEKSKQELNITASKPRDHVQVN 316

Qy 301 NFEALKTIONQREKIFAIBGTOTGSSSSPEHEMSQEGFAAITSNGPLLSITVGSYDWAG 360  
 Db 317 NFEALKTIONQREKIFAIBGTOTGSSSSPEHEMSQEGFAAITSNGPLLSITVGSYDWAG 376  
 Qy 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPRYQHILGLVAMFR 420  
 Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNRVQSLVGLGAPRYQHILGLVAMFR 436  
 Qy 421 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVSCPL 480  
 Db 437 QNTGMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYEYQTRGGQSVSCPL 496  
 Qy 481 PRGORARWQCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 540  
 Db 497 PRGORARWQCDAVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYLYF 556  
 Qy 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLDLTVGAQGHVLLLRSQ 600  
 Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVLDLTVGAQGHVLLLRSQ 616  
 Qy 601 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRCVCHVOKSTRDLREGOIQSVVT 660  
 Db 617 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRCVCHVOKSTRDLREGOIQSVVT 676  
 Qy 661 YDLALDSGRPHSRAVFNKSTRROTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 720  
 Db 677 YDLALDSGRPHSRAVFNKSTRROTQVLGLTQTCTETLKLQLPNCIEDPVSPIVLRNF 736  
 Qy 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPFEKXCGNDNICQDDLSITFSFMSLDCLVVG 780  
 Db 737 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPFEKXCGNDNICQDDLSITFSFMSLDCLVVG 796  
 Qy 781 GPREFNVTVTRNDGDSYRTQVTFPEPLDLSYRKVSTLONORSORSWRLACESASSTEV 840  
 Db 797 GPREFNVTVTRNDGDSYRTQVTFPEPLDLSYRKVSTLONORSORSWRLACESASSTEV 856  
 Qy 841 SGALKSTSCSINHIPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900  
 Db 857 SGALKSTSCSINHIPIFENSEVFNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916  
 Qy 901 QLELPVKYAVYVMTSHGVSTKYILNFTASNTSRVMOHQVSNLQORSIPISLVFLVPV 960  
 Db 917 QLELPVKYAVYVMTSHGVSTKYILNFTASNTSRVMOHQVSNLQORSIPISLVFLVPV 976  
 Qy 961 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVNCIAVCQRIQCDIP 1020  
 Db 977 RLNQTVLWDRPQVTFSENLSSTCHTKERLPSSHDSFLAELRKAPVNVNCIAVCQRIQCDIP 1036  
 Qy 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1080  
 Db 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVFTLLPGQAFVRSQTET 1096  
 Qy 1081 KVEPFEPVNPPLIVGSSVGLLALLITAAIYKLGFFKQYKQDMSEGPPGAEQ 1137  
 Db 1097 KVEPFEPVNPPLIVGSSVGLLALLITAAIYKLGFFKQYKQDMSEGPPGAEQ 1153

## RESULT 2

S00551

leukocyte surface glycoprotein Mac-1 alpha chain precursor - mouse

N;Alternate names: complement-3 receptor alpha chain

C;Species: Mus musculus (house mouse)

C;Date: 30-Sep-1989 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004

C;Accession: S00551; 159078

R;Pyteila, R.

EMBO J. 7, 1371-1378, 1988

A;Title: Amino acid sequence of the murine Mac-1 alpha chain reveals homology with the

A;Reference number: S00551; MUID:88312584; PMID:3044779

A;Accession: S00551

A;Molecule type: DNA

A;Residues: 1-1153 &lt;PYT&gt;

A;Cross-references: UNIPROT:P05555; EMBL:X07640; NID:g52982; PIDN:CAA30479.1; PID:g52982

A;Note: the authors translated the codon CAC for residue 569 as Gln

R;Sastre, L.; Roman, J.M.; Teplow, D.B.; Dreyer, W.J.; Gee, C.E.; Larson, R.S.; Roberts, Proc. Natl. Acad. Sci. U.S.A. 83, 5644-5648, 1986  
A:Title: A partial genomic DNA clone for the alpha subunit of the mouse complement recep  
A:Reference number: 159078; MUID:86287312; PMID:12942940  
A:Accession: 159078  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 11-44 <RES>  
A:Cross-references: GB:M14293; NID:g198993; PIDN:AAA39484.1; PID:g554193  
C:Genetics:  
A:Gene: Mac-1  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C:Keywords: cell adhesion; glycoprotein; transmembrane protein  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-1153/Product: leukocyte surface glycoprotein Mac-1 alpha chain #status experimental  
F:148-318/Domain: von Willebrand factor type A repeat homology <WMA2>  
F:1106-1129/Domain: transmembrane #status predicted <TM>

Query Match 76.2%; Score 4476; DB 2; Length 1153;  
Best Local Similarity 74.3%; Pred. No. 6.1e-300;  
Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;

QY 1 FNLDTENAMTFOENARFGSGVOLQGRVVGAPQEIIVANQSGISYQCDYSTGSCPEI 60  
Db 17 FNLDTEHPMTFOENAKFGQNVVLGGTSVVAAPQBAKAVNOTGALYQCDYSTSRCHPI 76

QY 61 RLQVPVEAVNMSGLSLAATTSPPOLACGPTVHOTCSENTYVYKGLCFLEGSNLRQPOK 120  
Db 77 PLQVPPEAVNMSGLSLAVSTVPOQLLACGPTVHONCKENTYVNGLCYLFSGNLLRPPQ 136

QY 121 FPALRGCPQEDSDIAPLIDGSGSIIIPDFRRMKEFVSTVMEOLKSKTILFSLMQXSEEF 180  
Db 137 FPEALRECPOESDIPLIDGSGSINNIDFQKMKFVSTVMEQFKSKTILFSLMQXSEEF 196

QY 181 RIHTFKBFQNNENRSLVKPIITQLGRTHATGVRKVRIBELNITNGARKNAFKILIVI 240  
Db 197 RIHTFENDFRNFSRSHVSPKQLNGRTKTASGIRKVRVRELFHKINGARENAKILIVI 256

QY 241 TDEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKROBLNTIASPPRDHVPQVN 300  
Db 257 TDEKFGDPLGYEDVPEADREGVIRVVGDAFRSEKROBLNTIASPPRDHVPQVN 316

QY 301 NFALKTIQNLREKIPAEIGTQTGSSSPSEHMSQEGFSAALITNGPLSLTVGSVDWAG 360  
Db 317 NFEALNTIQNLQEKFAIGTQTGSSSPSEHMSQEGFSAALITNGPLSLTVGSVDWAG 376

QY 361 GVFLYTSKEKSTPINTRVDSMDNDAYLVGAAILRNVRQSVILGAPRYOHIGLVAMFR 420  
Db 377 GAFLYTSKDKVTFINTTRVDSMDNDAYLVGASAVILRNVRQSVILGAPRYOHIGLVAMFR 436

QY 421 QNTGMESNANVKGTQIGAVFGASLCSVDVDSNGSTDLVLIGAPHYYEOTRGQVSVCP 480  
Db 437 ENFGTWEPTSIAGSQIGSTYFGASLCSVDMDAGNTNLIIGAPHYYEOTRGQVSVCP 496

QY 481 PRQGRARWQCDAYLYEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNKGAVYLF 540  
Db 497 PRG-RARWQCEALLHGDQGHWPGRFGAALTVLGDVNGDKLTDVAIGAPGEQNGQAVYIF 555

QY 541 HGTSGSGISPSHSQRIAGSKLSRPLQYFGOSLGGQDLTMDGLVDITVGAQGHVLLRSQ 600  
Db 556 YGASIASLSASHSHRIIGAHFSPGLQYFGOSLGGQDLTMDGLMDLAVGAQGHVLLRAQ 615

QY 601 PVLVRVAIMEFNEPREVARNVECDQVKGKEAGEVRVCLHVOKSTRDLRREGIQISVVT 660  
Db 616 PVLRLSATMEFSPKKVARSVFACQEQVLKNKDAGEVRVCLVRVKNTRDLRREGIQISTVT 675

QY 661 YDLALDSGRPHSAFVNETKNSRRTOTVGLGLTQTCETLKLQPCNIEDPVSPVILRLNF 720  
Db 676 YDLALDPVSRIRAPFDETKNTRTOTVGLGLMOKETLKLILPDCVDVSPVILRLNY 735

QY 721 SLVGTPLSAFGLRPVLAEDAQRLLFTALFPFEXKNCNDNICDDLSITFSMSLCLVVG 780  
Db 736 TLVGEPLRSGNLRPVLAMDQARFFATMPFEXKNCNDNICDDLSITFSMSLCLVVG 795

QY 781 GPREFNVTVRNDGEDSYQTQVTFPPPLDLSYVKVSTLQNRQSRWRL-ACSSASSTE 839  
Db 796 GPQDFNMSVTLRNDGEDSYQTQVTFPPPLDLSYVKVSTLQNRQSRWRL-ACSSASSTE 855

QY 840 VSGALKSTSGSINHPIPPENSEVENTITFDVDSKASIGNKLLKANVTSENMMPTNKTE 899  
Db 856 GHGALKSTTWNINPIPPANSEVTFNFDVDSHAFSGNKLKLLKAI VASENNMERTHKT 915

QY 900 FQLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVQMHOVQVSNLQSRSLPISLVFLVP 959  
Db 916 FQLELPVKYALYIMVTSDESSIRYLNFTASEMTSKVQHQQYQFNLLQSRSLPVSVVFWIP 975

QY 960 VRLNQTVINDRPOQVTSSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQDI 1019  
Db 976 VQINVTVDHPQVIFESQNLSSACHTEQKSPHNSFRDQLERTPVLNCSVAVCKRIQCDL 1035

QY 1020 PFTGIOEFNATLKNLSFDWYIKTSHNHLIIVSTABILFENDSVFTLLPGGAFVRSOTE 1079  
Db 1036 PSFTQEIFNVTILKNLSFDWYIKTSHGHLIIVSSTEILFENDSAFALLPQESTVRSKTE 1095

QY 1080 TKVEPEVNPPLPLIVGSSVGGLLALLIITAAALYKLGFFKRYQKDMSEGPPGABPQ 1137  
Db 1096 TKVEPEVNPPLIVGSSIGGLVALLIITAGLYKLGFFKRYQKDMNEAAPQDAPQ 1153

RESULT 3  
RWHLIC  
cell surface glycoprotein CD11c precursor - human  
N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain  
C:Species: Homo sapiens (man)  
C>Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: A36584; A35543; S00864  
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
J. Biol. Chem. 265, 12750-12751, 1990  
A:Reference number: A36584  
A:Contents: erratum  
A:Accession: A36584  
A:Molecule type: DNA  
A:Residues: 1-1163 <COR>  
A:Cross-references: UNIPROT:P20702  
A>Note: this revision to the sequence from reference A35543 includes the carboxyl end  
R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.  
J. Biol. Chem. 265, 2782-2788, 1990  
A:Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule  
A:Reference number: A35543; MUID:90153906; PMID:2303426  
A:Accession: A35543  
A:Molecule type: DNA  
A:Residues: 1-834 <CO2>  
A>Note: this sequence has been revised in reference A36584  
R:Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.  
EMBO J. 6, 4023-4028, 1987  
A:Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte  
A:Reference number: S00864; MUID:88166645; PMID:3327687  
A:Accession: S00864  
A:Molecule type: mRNA  
A:Residues: 1-755, 'U', 757-1163 <CO3>  
A:Cross-references: GB:M81695; EMBL:Y00093; NID:g487829; PIDN:AAA59180.1; PID:g487830  
A>Note: part of this sequence was confirmed by protein sequencing  
C:Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on m  
C:Genetics:  
A:Gene: GDB:ITGAX; CD11C  
A:Cross-references: GDB:119758; OMIM:151510  
A:Map position: 16p11.2-16p11.2  
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
C:Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-1163/Product: cell surface glycoprotein CD11c #status predicted <MAT>  
F:20-1107/Domain: extracellular #status predicted <EXT>  
F:149-319/Domain: von Willebrand factor type A repeat homology <WMA4>  
F:1108-1133/Domain: transmembrane #status predicted <TM>  
F:1134-1163/Domain: intracellular #status predicted <INT>  
F:61,89,392,697,735,899,939,1050/Binding site: carbohydrate (Asn) (covalent) #status pre



Db 141 LQGRPGQECIKG-----NVDLVFLFDGSMISLPDEFOKILDFMCKDVMKLSNTSYQFAAV 196  
QY 175 QYSEEFRIHTEFEFQNNPNSLVKPIITQLLGRTHATGVRKVIRELLINITGARKNAF 234  
Db 197 QFSTSYKTEFDFDYVKWKDPDALLKRVKEMLLTNTFGAINVYVATEVFEELGAPDAT 256  
QY 235 KILIVITDGKFGDPLGYEDVPEADREGVIRVIGVGDAFRSEKSRQELNLTASKPPRD 294  
Db 257 KVLIIITDGE--ATDSNGIDAAND-----IIRVIGIGKHFQYKESQETLHKFASKPASE 309  
QY 295 HVFQNNFEALKITQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAITNGPELLSTVG 354  
Db 310 FVKILDTFEKLDLFTLQKIVIEGTSKDLTSFNMLSSSGISADLSRGHVAVGVAVG 369  
QY 355 SYDWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGAAA-IILNRVQSLVLGAPYQH 412  
Db 370 AKDWAGGFLDKADLODDTTFIGNEPTTPEVRAGYLGTVTWLPSRQKTSLSAGAPRYQH 429  
QY 413 IGLVAMFR--QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEOT 470  
Db 430 MGRVLLFQEPQGGHMSQVQTHGTQIGSYFGGELCGVDVDDGETELLILIGAPLFYGEQ 489  
QY 471 RGGQVSVCLPRQORARWQDAV--LYGEOQOPWGRFGAALTVLGDVNGDKLTVAGAP 528  
Db 490 RGRGVFIY-----QRRLQGFEEVSELQGDGPYPLGRFGEAITALTIDINGDLVDVAVGAP 544  
QY 529 GEEDNRCAYVLPHGTSGSLSPSHSQRISAGSKLSPLOYFGQSLSGGQDITMDGLVDLTV 588  
Db 545 LEE--QNAVYIFNGRHG-GLSPQSPQRIGTQVLSGIQMFGRSHGVKDLGGDLADVAV 601  
QY 589 GAQGHVLLRSQPVLRVKALMEFNPREVARNVPECDNVV-KGKEAGEVRVCLHVOKSTR 647  
Db 602 GABSQMLVLSRRPVDVMTLMSFPAEIPVHEVECSYSTSNKMKEGVNIITICQI-KSLY 660  
QY 648 DLREGOISWVYTDIALDSGRPHSRVAFNETKXSTRQVQLGLTQTCTBLKQLPNCI 707  
Db 661 PQF-QGRVANLYTYTLQDGHRTRRGLFPGGRHLLRNIAVT-TSMSCDTPSFHPVCV 718  
QY 708 EDPVSPVILRNESL---VGTPLS--AFGN-----LRPULAEADQLFTALPFPEKNCGN 757  
Db 719 QDLISPINVSLNFSLWEEGTTPRQDQAQGDIPPIILRPSLHSETWEI-----PFENKCGE 773  
QY 758 DNICQDDLSITTFMSLDCLVCGPREFNVTVVRNDGEDSYRTQVTFEPFLDLSVRKYS 817  
Db 774 DKCEANLVSPSPARSRLRLTAFASLSVELSLNLEEDAYVWQDLHPPGLSPKVE 833  
QY 818 TLQNRQSQRWRLACES--ASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKAS 875  
Db 834 ML---KPHSQIPVSCPELPEESRLSAL---SCNVSSPIFKAGHSVALQMMENTLVNS 887  
QY 876 LGNKLKLLKANVTSENN-----MPKNTKTEFQLELPVKYAVYVMVTVSHGVSTKYLNFASEN 931  
Db 888 WGDVSELHANVTCCNEDSLLEDNSATTI---IPILYPINILQDQEDSTLYVSFTPKGP 944  
QY 932 TSRVMOHQYV---SNLQORSIP-LSILVLPVRLNQTWDRPQVTFSENLSSTCHTK- 986  
Db 945 KIHQVKHMVQVRLQPSIHHDNITPLEAVGVQPPSEGIETHOWSVQMEPPV--PCHYED 1002  
QY 987 -BRLPSHSD--FLAELURKAPVNCISIAVCQRIQCIPFFGIOBEFNATLKNLSFDWYIK 1043  
Db 1003 LERLPDAAEPCLPGALFRCPVW-----FRQEILVQVIGTLELVEITE 1044  
QY 1044 TSHNHLIIVSTABILNDSVFTLLPGQAFVRSQETKVEPEFVNPFLPLIVGSSVGGLL 1103  
Db 1045 AS-SMFSLSSLSISNSKHFLYGSNASL-AQVVMKVDDVYVKQMLYLYLVSIGIGLL 1102  
QY 1104 LLALITAAALYKLFPPKRVQKMMSEG-GPPGAP 1136  
Db 1103 LLLLIIVLYKGVFFRNILKERKEAGRGVNGIP 1136

RESULT 5  
I56126

lymphocyte fuction-associated molecule-1-alpha - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 09-Jul-2004  
C;Accession: I56126  
R;Kaufmann, Y.; Tseng, E.; Springer, T.A.  
J. Immunol. 147, 369-374, 1991  
A;Title: Cloning of the murine lymphocyte function-associated molecule-1 alpha-subunit  
A;Reference number: I56126; MUID:91268576; PMID:2051027  
A;Accession: I56126  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: mRNA  
A;Residues: 1-1163 <RES>  
A;Cross-references: UNIPROT:P24063; GB:M60778; NID:9198785; PIDN:AAA39426.1; PID:919878  
C;Genetics:  
A;Gene: LFA-1  
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom  
F;151-315/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match 26.1%; Score 1533.5; DB 2; Length 1163;  
Best Local Similarity 34.0%; Pred. No. 4.6e-97;  
Matches 398; Conservative 217; Mismatches 458; Indels 99; Gaps 36;

QY 1 FNLDTENAMTFOENA-RGFGQSVVVLQSGSRVVVVGAPQEIIVAAQNRGLSYQCDYSTGCEP 59  
Db 24 YNLDTPTQSFQAQGRHFGYQLIEDG-VVVGAPGE---GDNTGGLYHCRITSSFECP 79  
QY 60 IRLQVPEAVNMSLGLSLAATTSPQLLAGPTVHOTCSENTYVYKGLCFGLFSGNLRQPPQ 119  
Db 80 VSLH-GSNHTSKYLGMTLATDAAKGSLACDPLGSLTCDQNTYLSGLCYLFPQSLGPGML 138  
QY 120 KPEALRGCPQEDSDIAFLIDSGSIIIPHDPRMKEFVSTVMEQLKKSKTLFSLMQYSBE 179  
Db 139 QNRPAQECMKGVKDLVFLFDGSQLDRKDFKILEFMKVMRKLSTNTSYQFAAVQFSTD 198  
QY 180 FRHFFTEKF-QNNPNRSLVVKPIITOLLGRTHATGVRKVIRELLINITGARKNAFKILI 238  
Db 199 CRTEFTFLDYVKQKNPDLVLSGVQPMFLTTNTFRAINVVAHVFFKEGSGARDATKVLV 258  
QY 239 VITDGSKFGDPLGYEDVPEADREG-----VIRYVIGVGDAFRSEKSRQELNLTASKP 291  
Db 259 IITDG-----EASDKGNISAAHDITRYIIGKHFVSVQKQTLHIFASEP 304  
QY 292 PRDHVQVNNFEALKITQNLREKIFAIEGTQTGSSSSPEHEMSQEGFSAITNGPELLS 351  
Db 305 VBEFVKILDTFEKLDLFTDQRIYAIIEGTNRQDLTSFNMLSSSGISADLSRGHVAV 364  
QY 352 TVGSDVWAGGVF-LYTSKEKSTFINMTRVDSMDNDAYLGAAA-IILNRVQSLVLGAPR 409  
Db 365 AVGAKDWAGGFLDLREDLQGAIFVQGEPLTSDVRCGYLGYTVNWTSSRSLPILAGAPR 424  
QY 410 YOHIGLVAMFR--QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYY 467  
Db 425 YQHVQVILLFQAPEAGGRWNQTKIEGTQIGSYFGGELCSVDLDQDGEAELLIGAPLFF 484  
QY 468 EQTRGQSVSVCLPRQORARWQDAVLYGEOQOPWGRFGAALTVLGDVNGDKLTVAGIA 527  
Db 485 GBQRGGRVTVY---QRRQSLFEMVSELQGDGPYPLGRFGAATALTIDINGDLTIDVAVGA 541  
QY 528 PGEENRGAAYVLPHGTSGSGISPSHSQRISAGSKLSPLOYFGQSLSGGQDITMDGLVDLT 587  
Db 542 PLEE--QNAVYIFNGKPG-GLSPQSPQRIGQAVFGFIRWFGERSIHGVKDLGGDLRADVV 598  
QY 588 VQAQGHVLLRSQPVLRVKALMEFNPREVARNVPECDNVVQVVKGEAG-EVRVCLHVQKST 646  
Db 599 VGAERGVVVLSSRPVVDVVTLSFSPEEIPVHEVECSYAREBEQKHGVKLCACFRKPLT 658  
QY 647 RDRLEGQIQSVTVYDLDLDSGRPHSRVAFNETKSTRQVQLGLTQTCTETKLQLPNC 706  
Db 659 PQ--FQGRLLANLSYTLQDGHMRMSRGLFPDGSHELSTNTSITP-DKSCLDLFFHFFPIC 715  
QY 707 IEDPVSPVILRNFSLV---GTPLSAFGN-LRPVLAEADAQLFTALFPPEKNCNDNICQ 762  
Db 716 IQDLISPINVSLNFSLEEGTTPRQDQKRAMQPIILRPSIHTV-TKEIPKKNCKEDKKCE 774



```

QY 763 DLSITFSFMSLCLVVGPP-----REFNVTVRNDGDSYRTQVTFEPFLDLSYRKV 816
Db 775 ANLITSSPARK-----GPIRLMSASLAVETWLSNGSDAYVWRJLDLPFGRGLSPRKV 827
QY 817 STLQNORSQRWRLACSSASTVSGAL-KSTSCSINHPIPPENSRYTFNITFDVDSKAS 875
Db 828 EMLQ---PHSRMPVSCBEL--TEGSLTLTKLKNVSSPIFKAQEVSQVLMWFNTLNS 882
QY 876 LGNKLKANKVTSN-NMPTNKTEFQLELPVKYAVVMVTVSHGVSKYLNFTASENTR 934
Db 883 WEDFVELNGTVHCENENSSLOEDNSAATHIPVLPVNLTKQENENSTLYTSFTPKGPKTQ 942
QY 935 VMOHOYOVNLSQSLPISLVFLVPLVRLNQTVIWRDQ---VTFSENLS---TCHTK 986
Db 943 QVQHVIVQ---RIQSAVDHNMPT-LEALGVPRHSEDLITTVTSVQDPLVTVCHSE 996
QY 987 E-RUPSHSHFLAELRKAPVNCSTAVCQRIQCIPFFGIQEEFNATLKGNSLFDWIYKTS 1045
Db 997 DLKRPSE---AEQPCPLGV-----QFRCPIVF--RWEILIQVTGVLSKEIKAS 1042
QY 1046 HNHLIIVSTAEILLFNDVFTLLPQGAFAVRSOTETKVEPEVFNPLPLIVGSSVGGILL 1105
Db 1043 -STLSCLSSVSFNSSKHPHLYGSKA-SEAQVLVKVDLIHEKMLHVYVLSGGIGLVLL 1100
QY 1106 ALITTAALYKLGFFKRYQKDMM-SEGPPGARP 1136
Db 1101 FLIFALYKVGFFRNLRKEMADGGVNGSP 1132

RESULT 6
A53213
integrin alpha-E chain - human
C:Species: Homo sapiens (man)
C>Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 09-Jul-2004
C:Accession: A53213
R:Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A:Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (E) subunit. Un
A:Reference number: A53213; MUID:94164962; PMID:8119947
A:Accession: A53213
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1179 <SHA>
A:Cross-references: UNIPROT:P38570; GB:L25851; NID:g457244; PID:g457245
C:Genetics:
A:Gene: ITGAE
A:Cross-references: GDB:330801
A:Map position: 17p13
C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
P:199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 19.6%; Score 1149; DB 2; Length 1179;
Best Local Similarity 29.1%; Pred. No. 1.5e-70;
Matches 343; Conservative 211; Mismatches 451; Indels 172; Gaps 39;

QY 45 GSLVQCDYSTGS--CEPI-RLQVP-----VEAVNMSLGLSLAATTSPQLLACGPTVHQ 95
Db 65 GPLHRCSLVQDEILCHPEVHVPIPKGRHGRGVTVVRSHHGVLICI-----QVLVRRP--HS 117
QY 96 TCSENTYVKGCLFLFGSNLRQPPQ-----SDIAFLDGSGLIIPDPRMKEFVSTWNEQL-- 119
Db 118 LSSELT---GTCSLGLPDLRPAQANFFDLENLLDPAQVDTGDCYSNKEGGEDDVNTA 174
QY 120 KFPALRGCPQED-----SDIAFLDGSGLIIPDPRMKEFVSTWNEQL-- 164
Db 175 QRRALEKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEEEDKEE 234
QY 165 KKSRTLFLSMOYSEEFTHFTFKFONPNPRSLVKPITQLLGRTHATATGVKRVIRELLN 224
Db 235 KCFECNFALVQGVQVQTEFLDRSDQVWASLARVQNTQVGSVTKTASAMQHLDLSFT 294
QY 225 ITGARKNAFKILIVITDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKROEL 284

```

## RESULT 7

A45226

integrin alpha-1 chain - human (fragment)

C:Species: Homo sapiens (man)

C&gt;Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

C:Accession: A45226

R:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.

J. Biol. Chem. 268, 2989-2996, 1993

A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit

```

Db 295 SSHGSRKAKSVAVVLTDGGIFEDPLNLTVINSPKMQGVREFAIGVGEFFKSARTAREL 354
QY 285 NTTASRPPRHVQVANNFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFSAAIT 344
Db 355 NLIASPDDETHAFKVTNYMALDGLLSKLRYNIIISMEGT---VGDALHYQAQIGFSAQL 411
QY 345 SNGP-LLSTVGSVDWAGGVFLY--TSKEKSTFINNTRVDSMDNDA---YLGYSAAAILRN 398
Db 412 DERQVLGAVGAFDWSGGALLYDTRSRGRFLNQTAATAAADAQAQYSYLGYAVVLHKT 471
QY 399 RVOSLVLGAPRYQHIGLVAMFR-QNTGMWESNANV-KGTQIGAYFGASLCSVDVDSNGST 456
Db 472 CSLSYVAGAPQYKHG--AVFELQKEGREASFPLVLEGEQMGSYFGSELCPVDIDMDGST 529
QY 457 DLVLIGAPHYYEQTRGQOVSVCPPLQCORARWQCDVLYGEOGPWGRFGAALTVLGDVN 516
Db 530 DFLVAAPFFVHVHGEGRVVYRLSE-ODGSFSLARLSLHGHGFTNARFGFAMAAGDLS 588
QY 517 GDKLTDVAIGAP---GEEDNR--GAVLPHGTSGSGISPSHSQRIAGSKLSPLQYFGQ 570
Db 589 QDKLTDVAIGAPLEGFGADGASFGSVIYNG-HWDGLSASPSQIRIRASTVAPGLQYFGM 647
QY 571 SLSGQDLTWMDGLVLTGVAQGHVLLRLSOPVLRVKAIMFNPREFVARNVFECDNVKVG 630
Db 648 SMAGGFDISGDGLADITVGTLGQAVVPRSPVRLKVSMAFTPSALP-----IGF 697
QY 631 KEAGEVRVCLHVQKSTDRRLREGIQSVTVYDLALDSGRPHSRVAFNETKNSTERTQVL 690
Db 698 NGVVNVELCFEI--SSVTTASESGLREALNFTLDDVDGQKRRRLQCDSDVSCSLCLREWS 756
QY 691 GLTOTCETLKLQLPN---CIEDPVSPIVLRNLSVLGTPLSAFGRNLRPVLAEQAORLFT 746
Db 757 SGOQLCEDL-LMPTEGELCEEDCFNSASVKVSQQL-OTPEGQTDHPQILDRTYEFAL 814
QY 747 ALPFFPKNCNDNICQDDLSTTFMSLDCLVGGGPREFNVTVTVRNDGDSYRTQVTF 806
Db 815 FQLPYEKACKNKLFCVAELQLA-TTVSQQLVGLTKELTLINLTNSGDSYMTSMALN 873
QY 807 FPLDLVRKYSTLQONQORSQSWRLACSSASTEVSGALKSTCSINHPIPEENSEVTFNI 866
Db 874 YPRNLQ-----LKMQKPPSPNIQCDDPPQV---ASVLINMNCRIGHFVL-KRSSAHVSV 923
QY 867 TFDVDSKASLGNKLLKANVTSENN---MPRTNKTEFQ---LELPKYAVYVMVTVSHGV 919
Db 924 VWOLEENAFNRTADIIVTVINSNERSLANEHTLQFRHGFAVLKSPSIMYVNTQGL 983
QY 920 S--TKYLNFTASENTSRVMQHQVSNLQORSPLFISLVFLVPLVRLNOTVIMDRPQVTFSE 977
Db 984 SHHKEFLFHVHGEN--LFGAEYQ-----LQICVPTKLRLGLQVAAVKKLRTQ 1028
QY 978 NLSSTCHTKERLPSHSDFLAELRKAPVYVNGSIACVQRIQCDDIPFFGIQEEFNATLKGNSL 1037
Db 1029 ASTVCTVTSQSORACAYSS-VQHVSEHWSVSCVIA-----SDKENVTVAARIS 1073
QY 1038 FDWYIKTSHNLLIVST-----AEILFNDVSFTLLPQGAFAVRSOTETKVEPEVFNPL 1091
Db 1074 WD-----HSELLKQVTEQLQILGEISFNKSLYGLNAENH--RTKLTIVFLKDEKXHS 1125
QY 1092 PLIVGSSVGGILLALITALYKLGFFKRYQKDMSE 1128
Db 1126 PIILKSGVGLLIVLILVILFKCGFFKRYQQLNLE 1162

```

A;Reference number: A45226; MUID:93155124; PMID:8428973

A;Accession: A45226

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-1151 <BRI>

A;Cross-references: UNIPROT:P56199

A;Experimental source: hepatoblastoma cell line HepG2

A;Note: sequence extracted from NCBI backbone (NCBIP:124326)

F;142-317/Domain: von Willebrand factor type A repeat homology <WVAL>

```
Query Match      18.8%; Score 1102.5; DB 2; Length 1151;
Best Local Similarity 27.8%; Pred. No. 2.4e-67;
Matches 344; Conservative 211; Mismatches 487; Indels 195; Gaps 44;

QY 1 FNLDTENAMTFQENARG-FCQSVVQL---QGRVVVVGAPQEIIVAANQRGSLYQCDISTGS 56
Db 1 FNVVNSMTFSGVEDMFGYVVOQYENBEGKVLGSLVGPQPKNTGVDVYKCPVGGRGE 60
QY 57 CEP-IRLOQVPEA-----VNMSLGLSLAATTSPQLLAGCPTVHQTCSENTVYKGL 106
Db 61 SLPCVKLDLPVNTSIPNVTVBKENTMTFGSTL-VTPNPGGFLACGLYAVRCGHLHYTGI 119
QY 107 CFLFGSNLRQOPKPFPEARLGCQEDSDIAFLDGGSIIPHDFRMRKPFVSTVMEOLK- 165
Db 120 CSDVSPTFQVNSIAP--VOECSTQ-LDVIIVLDGNSIYFWD--SVTAFNLDLKRMDI 174
QY 166 -KSKTLFSLMOYSEEFRIHFTKFEFQNNPNRSLVKPITOLLGR-THTATGVRKVIRELL 223
Db 175 GPKQTQGVIGQVGENVTHEFNLNKYSTEEVLVAKKIVQGRQMTALGTDRTARKEAF 234
QY 224 NITNGARKNAFKILIVITDGEKFGDPLGVEDVPEADREGVIRYVIGVDADR-SE 278
Db 235 TEARGARRGVKVMVITDGEH-DNHRLLKVIQDCEDENIQFSAIILGYSNRGNLSTE 293
QY 279 KSKQELNTASKPRPHVFOVNNPEALKITQNLREKIPAEIQTGTGSSSPFHEMSQEG 338
Db 294 KFVEETKSIASEPTKHFNFVNSDELAVTIVTKLGERIFALEATADQSAASFEMESQTG 353
QY 339 FSAAITSNGLPLSTGVSYDWAGGVFLYTSKE-----KSTF-INMTRVDSMDNDAYLGAA 392
Db 354 FSAHYSDVMVLGAVGAYDWNGTVVMQKASQIIIPRTTFNVESTKKNELP-ASYLGTV 412
QY 393 AILNRVQSL-VLGPARYOHIGLVAMFRONTQGMESNANVKGTQICAYFGASLCSVDVD 451
Db 413 NGATASSGDVLYTAGQPRYNTQVLIYRMEDGNIKILQTLSEQIGSVFSLITLTDID 472
QY 452 SNGSTDVLILGAPHY-----YEQTR-GGQVSVCLPRGQARWQCDAVLY 495
Db 473 KDSNTDILLVGPAMYMGTEKEBQGVYVYALNQTRFQYQMSLEPIKQTCSSRQHNSCTT 532
QY 496 GQCGQPWG-RFGAALTVLGDNVNDKLTDAIAGAPGBEDNRGAVLYLFGTSGSGISPSHQ 554
Db 533 ENKNFPGARFGTAIAVKDLNLGDNFNDIVIGAPLEDHGGAVIYTHG-SGKTIRKEYAQ 591
QY 555 RIAGSKLSPRLQYFGQSLSGGQDLTMDGLVLDLTGVAQGHVLLRLSPVLRVKAIMEFNR 614
Db 592 RIPSQGDGKTLKFFQGSIHGEMDLNGDGLTDTVIGLGAALFWSRDAVAVKVTWVNFEN 651
QY 615 EVARNVFCNDQVVKKEAG--EVRVCLHVQ-KSTDRDLREGIOISVTVYDIALDSGRPH 671
Db 652 KVNIOKKNCH--MEGETVCINATVCFVKLSKEDTIVEADLQ-----YRVLDSLRLQI 704
QY 726 SRAVFNET-----KNSTRQQTQVLGLTQTCETKLQLPNCIEDPVPVPLRLNLSLVT 725
Db 705 SRSFSGTQERKQVNIYRKSEB-----TKHSFYMLDKHDFQDSVR-----ITLDFNLT-D 755
QY 726 PLSAFONLRPVLAEDAQRLLTALFPPEKNCNDNIQDDLSITTFSPMSLDCLVVGGPRE- 784
Db 756 PENG-----PVLDDSLPNSVHEIYIPAKDCGNKEKICISLHLVATTEKDLILVRSQNDK 810
QY 785 FNVTVTVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRSRWLUACASASTEVSGAL 844
Db 811 FNVSLTVKNTKDSAVNTRTIVHYSPLNVFSGIEAIQKD-----SCSEN----- 853
```

```
QY 845 KSTSCSINHPIFBNSEVTEFNITFDVDSKASLGN-KULLKANVTSENMPRTNKTEFOLE 903
Db 854 HNTCKVGYBFLRRGEMVTFKILFQFNTSYPMENVTIYLSATSDSEBPPETLSDNVVIS 913
QY 904 LPVKYAVVMVVTSGHGVSTKYLNFTASNTSRVMQHOYQVSN-----LGQRS-----L 950
Db 914 IPVKYEVGLQFYS-SASEYHISIAANETVPEVINSTEDIGNEINIFYLIRKSGSFPMPLE 972
QY 951 PISLVF-----LVPVRLNQTVIWRDPQVTFSENLSSTCTHKE-----RLPS 991
Db 973 KLSISFPMNTSGVPLVPTGLSS-----SENANCRPHIFEDPFSINSKQWTT 1021
QY 992 HSDFLAELRKAPVNVGSIACVQRIQCDIPFGIOE-----EFNATLK 1033
Db 1022 STD---HLRGKGLDNCNTCKFATITCNLTSSDISQVNVSLILWKPTIKSYFSSNLNLTIR 1078
QY 1034 GNLSFDWYIKTSNHLILVSTABEILFNDSVFTLLPGOGAFVRVRSQTETKVEPFEVNPPLP 1093
Db 1079 GEL-----RSENASLIVSSN-----OKRELAIQISKQGLPGRVPL 1114
QY 1094 --IVGSSVGGLLLLALITAAALYKLGFFKQYKDMSE 1128
Db 1115 WVILLSAFAGLLMLLILALWKIGFFKPLKKKMEK 1151

RESULT 8
I45914
Integrin alpha 2 subunit - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C;Accession: I45914
R;Ramata, I.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A;Title: Identification of putative ligand binding sites within the I-domain of integrin
A;Reference number: A54402; MUID:94193647; PMID:7511592
A;Accession: I45914
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1170 <KAM>
A;Cross-references: UNIPROT:P53710; GB:I25886; NID:G439695; PIDN:AAB59255.1; PID:G439695
F;161-336/Domain: von Willebrand factor type A repeat homology <WMA2>
```

```
Query Match      18.5%; Score 1084; DB 2; Length 1170;
Best Local Similarity 27.7%; Pred. No. 4.6e-66;
Matches 336; Conservative 216; Mismatches 495; Indels 168; Gaps 47;

QY 1 FNLDTENAMTFQ-ENARGFGQSVVQL---QGRVVVVGAPQEIIVAANQRGSLYQCD--DYST 54
Db 19 YNVGLPKALIFSGPSSEQEGYAVQQFIPNKGWLLVGSFWSGFPKRMGDVYKCPVDLST 78
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPQLLAGCPTVHQTCSENTVYKGLC 107
Db 79 TTCEKLNLTSTSMGNSVNTKMTNMSLGLTLTRNVGTGGFLTCGLPLWAQCGSQYTTTGV 138
QY 108 FLFGSNLRQOPKPFPEARLGCQEDSDIAFLDGGSIIPHDFRMRKPFVSTVMEOLK-- 165
Db 139 SDVSPDF-QLRTSFAPAVQTCP-SFIDVVVVCDESNIYPWD--AVKNLEKFEVQGLDIG 194
QY 166 KSKTLFSLMOYSEEFRIHFTKFEFQNNPNRSLVKPITOLL-----GRTHATGVRKVIRE 221
Db 195 PKTQOMGLIQYANPRVFNLTNFKSD---EMIATSTQTFQYGGDLTNTFKAIQIARTD 251
QY 222 LNIITNGARKNAFKILIVITDGEKFGDPLGVEDVPEADREGVIRYVIGV-----GDAPR 276
Db 252 AYSTAAGRPGATKVMVVVTDGESH-DGSKLKAVIDQCNKONILRFGIAGVLYNENALD 310
QY 277 SEKSQELNTASKPRPHVFOVNNPEALKITQNLREKIPAEIQTGTGSSSPFHEMSQ 336
Db 311 TKNLKETKATASIPTEPHFFNVSDADLLEKAGTIGEQIFSGIEGTVOQ-GDNFQWMSQ 369
QY 337 EGFSAAIT--SNGPLSLTVGSYDWAGGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGVA 391
```

370 VGEAEVSPQNNILMLGAVGADWSGVVQKTPHGHLIFSKQAFQEQILQDRNHSSYLGS 429  
392 AAILNRVOSLVGAPRYOHIGLVAMFRONTGWESNANV-----KGTQIGAYPGASL 445  
430 VASISTGNSVHFVAGAPRANVTGQIVLYSVN-----ENGNVTIQSQRGDQIGSYFGSVL 484  
446 CSVDVDSNGSTDVLVIGAPHYEQTR--GGQVSVCLPRGORARWQCDVAVLGBQGPWG 503  
485 CAVDVNDKTTDVLVAGAPMYMDLKKBEGRVYLFTIKG-ILNWH--QFLEGNGLENA 541  
504 RFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLFHGTSGSGISPSHSQRLAGS--KL 561  
542 RFGSAIALSDINMDGENDVIVGSPLENQNSGAVIYNGHEGM-IRLRYSQKILGSDRAF 600  
562 SPLQYFGQSLGGQDLTMGLVDLTVGAQGHVLLRLSQPVLRVKAIMEPNPREVARNVF 621  
601 SSHLQYFGRSLDGYGLDNGSDITDVSVGAFQGVVQLMSQSIADVSVDASETPKKI--TL 658  
622 ECNDQVYKGEAGVRVCLHVQKSTRDLRREGQIQSVVTVDLALD---SGRPHSRAVEN 677  
659 NKNRAE-----KKLQCF---SAKPRPTNNQNVAVIVNITIDEQFSSRVLSRGLFK 707  
678 ETKNSTRRQTVGLGTQCE--TLKLQLPNCIEDPVSPIVLRNLSL--VGTPLSAFNL 733  
708 ENNERCLOKTMIVSQACRCSEYIITHIQEPS---DIISPLNLCMNISLENPCF----- 756  
734 RPVLAEADQRLFTALFFPERKCNMDNICQDDLSITF-----SFMGLDCLVGGPREFNVTV 789  
757 NPALAEAYSETVKFISIPFKDCGDDGVCISDLVLNVQOLPATQOQPFIVSNQNKRLTFSV 816  
790 TVRNDGDSYHTQVTFPFPLDLSYRKVSTLQNRQSRQSWRLACESAST-EVSGALKSTS 848  
817 QLNKKESAYNTEIVDPFSENLFF-----ASWSMPVDGTEVTQCIASQSKSVT 864  
849 CSINHPFIPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFFOLELPVKY 908  
865 CNVGYPALKSKQVTFITNFDNLQ-NLQNASISPRALSQSENMDNSVNLKLSILY 923  
909 AVYMWTVSHGVSTKYLVNTASENTSRVMQHOYQVSNLQOR-----SLPTSLFLV 958  
924 DAEIHT-RSTNINFEYSLDGNSSVV-HSFE--DIGPKFIFSIVTGTGSPVPSMA--- 976  
959 PVRLNQTVIMDRPQVPSFN--LSSTCHTKB-----RLPSHSDFLAE- 998  
977 -----SVIIHPQVTKDNPLMLYTGHTQAGDISCEABINPLKIGQTSSSVSFKSEN 1030  
999 LRKAPVNCSTAVCQRIQCDIPFFGIOBEFNATLKNLSFDWYIKTSHNHLIIVSTABI- 1057  
1031 FRHIKELNCRATASCNIMCWLRLDQVKGVEFLNVSTRINWGTFAASTFQTVQLTAAABID 1090  
1058 LFNDVSFTL-----LPCQGAFVRSQETKVEPPE-VPNPLPLIVGSSVGLLILALITA 1110  
1091 TYNPQIYVIEBNTVIP-----LTIMKPKHEKVEVPTGTVIVGSGVIAGILLALLVA 1140  
1111 ALYKLGFPPKROYKDM 1125  
1141 ILWLKGFPPKRYKYM 1155

## RESULT 9

S44142  
VLA-2 protein homolog - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S44142  
R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze  
submitted to the EMBL Data Library, January 1994  
A:Description: The mouse VLA-2 homologue supports collagen and laminin adhesion but not  
A:Reference number: S44142  
A:Accession: S44142  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1178 <EDE>

A;Cross-references: UNIPROT:Q62469; EMBL:Z29987; NID:g473098; PID:CAA82877.1; PID:g4747  
F;169-344/Domain: von Willebrand factor type A repeat homology <VMA2>

Query Match 18.2%; Score 1072; DB 2; Length 1178;  
Best Local Similarity 28.1%; Pred. No. 3.1e-65;  
Matches 344; Conservative 206; Mismatches 487; Indels 188; Gaps 44;

QY 1 ENLDTENAMTFQ-ENARGFGQSVVQL--OGSRVVVGAPQEIIVAANORGSLYQC--DYST 54  
DB 27 YNVLPGAKIFSGPSSSEQFGYSVQQLTNPQGNWLLVGSPPWSGPPENRMGDVYKCPVDLPT 86  
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATTSPQLLACGPTVHOTCSENTYVKGCL 107  
DB 87 ATCEKLNQNSASISNVTEIKTNMSLGLTLTRNPGTGGFLTCGPLWAHQCGNQYATGIC 146  
QY 108 FLFGSNLRQPOQ--KFPEALRCGPQEDSDIAFLIDGSGSIIHPDPRMKEFVSTVMEQL 164  
DB 147 ----SDVSPDFQFLTSPFAVQACPSL--VDVVVVVCDENSIYP--WEAVKNFLVKFTVGL 199  
QY 165 K--KSKTLFSLMOYSEBFRIHFTFKFQNNPNPRSLVKPITQLILG-RTHATGVRKVVIRE 221  
DB 200 DIGPKTKQVALIQVANEPRILINLNDPETKEDMVQATSETFQHGDLTNTFRAFEFARDY 259  
QY 222 LLNITNGARKNAFKILIVITDGEKFGDPLGYEDVIPADREGVIRYVIGV-----GDAFR 276  
DB 260 AYSQTSGRPGATKVMVVTVDGESH-DGSXLKTVIOQCNDDEILRFGI AVLGYLNRNALD 318  
QY 277 SEKSRQBLNTIASPPRDHVFQVNNFALKTIQNOLREKFIARIGTOTGSSSSEHEMSQ 336  
DB 319 TKNLKEIKAIATSTERYFNFVADEAALLKAGTLGEQIFSIETGVQ--GDNFQMEWAQ 377  
QY 337 EGPSA--AITSNGLPLSTVGSYDWAGVFLYTSKEKSTFINMT--RVSDSMN-DAYLGYA 391  
DB 378 VGFSAADYAPQNDILMLGAVGADWSGTLVQETSHKPVIFPKQAFDQVLQDRNHSSFLGS 437  
QY 392 AAILNRVOSLVGAPRYOHIGLVAMFRONTGWESNANV-----KGTQIGAYPGASLCS 447  
DB 438 VAAISTEDGVHFVAGAPRANVTGQIVLYSVN---QGNVTVIQSHRGDQIGSYFGSVLCS 494  
QY 448 VDVDNSGSTDLVLIGAPHYEQTR--GGQVSVCLPRGORARWQCDVAVLYGEOQGPWGRF 505  
DB 495 VDVKDKTITDVLVAGAPMYMDLKKBEKGVLFITIKGILNQHO---FLEGFTGTGNARF 551  
QY 506 GAALTVLGDVNGDKLTDVAIGAPGERDNRGAVYLFHGTSGSGISPSHSQRIAGSKLSPR- 564  
DB 552 GSAITAAALSDINMDGENDVIVGSPVENENSGAVIYNGHQGT-IRTKYSQKILGNGAFR 610  
QY 565 -LQYFGQSLGGQDLTMGLVDLTVGAQGHVLLRLSQPVLRVKAIMEPNPREVARNVPEC 623  
DB 611 HLQFFGRSLDGYGLDNGSDITDVSIGALGVQLWSQSIAVIAEALFTP----- 660  
QY 624 NDQVVKGEAGEVRVCLHVQKSTRDLRREGQIQSVVTVDLALD---SGRPHSRAVENET 679  
DB 661 -DKITLANKDAKITKLKCFRAEFAPAGQNNQV--AILFNMTLDADGHSRVRTSGVFREN 717  
QY 680 KNSTRROTQVGLTQTCET--LKLQLPNCIEDPVSPIVLRNLSFSLVGTPLSAFAGNLREPL 737  
DB 718 SERFLQKNMVNVEVQKCSHHISIQKPS---DVVNPLDLRVDISLENPGTS-----PAL 768  
QY 738 AEDAQRULTALFPPEKNGNDNICODDLISI-----TFSPMSLCLVVGGRFENVT 788  
DB 769 EAYSETVKVFSIPFYKECGSDGICITSDILVDVQOLPAIQTQSF-----IVSNQNKRLTFS 823  
QY 789 VTVNDGEDSVRTQVTFPFPLDLSYRKVSTLQNRQSRQSWRLACESAST-EVSGALKST 847  
DB 824 VILKNRGESANTVVLAEFSENLF-----ASFSMPVDGTEVTCEVGSQSKSV 871  
QY 848 SCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFFOLELP 905  
DB 872 TCDVGYPALKSEQQVTFITNFDNLQ-NLQNAALNFAQFSESQ--ETNKADNSVSLATIP 928  
QY 906 VKYAVYMWTVSHGVSTKYLVNTASENTSRVMQHOYQVSNLQOR-----SLPTSLV 955

```

Db 929 LLYDAELHILT-RSTNINFYEISSDENAPSVIK---SVEDIGPKFIFSLKVTAGSAPVSWA 984
QY 956 FLV-----PURLNQTIVDRPQVTF-SENLS 980
Db 985 LVTHIHPQYTKENKPLLYLTGTQDQAGDISCTABINPLKPHTA-----PSVSFKNENR 1040
QY 981 STCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIPFFGIGQBEFNATLKNLSFDW 1040
Db 1041 ---HTKE-----LDCRTTSCSNITCWLKDLHMKAEYFINVTRVWNR 1080
QY 1041 YKTSNHLIIVSTAELFNDVSFTLLPGGAFVRSQTETKVPFFVFNPLPLIVGSSVG 1100
Db 1081 FFASTFQTQVLTAAABIDTHNPOLFVEENAVTIPLMIMKPTKAEVPT--GVIIGSIIA 1138
QY 1101 GILLALLITAALYKLGFEKQYKDM 1125
Db 1139 GILLALLAMTAGLWKLGFEKQYKDM 1163

RESULT 10
A33998
N;Integrin alpha-2 chain precursor - human
N;Alternate names: CD49b; platelet glycoprotein GPIa; VLA-2/collagen receptor alpha-2 ch
C;Species: Homo sapiens (man)
C;Date: 30-Mar-1990 #sequence revision 18-Sep-1992 #text_change 09-Jul-2004
C;Accession: A33998; B56793; A53117
R;Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A;Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet
A;Reference number: A33998; MUID:89308879; PMID:2545729
A;Accession: A33998
A;Molecule type: mRNA
A;Residues: 1-1181 <TAK>
A;Cross-references: UNIPROT:P17301; GB:X17033; NID:933906; PIDN:CAA34894.1; PID:933907
A;Note: The authors translated the codon GAR for residue 802 as Gln, GTC for residue 803
R;Catmell, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A;Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIIb and
A;Reference number: A56793; MUID:92061944; PMID:1953640
A;Accession: B56793
A;Molecule type: protein
A;Residues: 30-43 <CAT>
A;Experimental source: platelet
R;Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A;Title: The human alpha-2 integrin gene promoter. Identification of positive and negati
A;Reference number: A53117; MUID:94103255; PMID:8276836
A;Accession: A53117
A;Molecule type: DNA
A;Residues: 1-16, 'V', 18-21 <ZUT>
A;Cross-references: GB:I24121; NID:9400342; PIDN:AAA16619.2; PID:94583535
A;Note: authors translated the codon GTA for residue 17 as Leu
C;Genetics:
A;Gene: GDB:ITGA2; CD49B
A;Cross-references: GDB:128031; OMIM:192974
A;Map position: 5q11.1-5q11.2
C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;30-1133/Domain: extracellular #status predicted <EXT>
F;172-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;1105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 18.2%; Score 1069; DB 2; Length 1181;
Best Local Similarity 27.0%; Pred. No. 5e-65;
Matches 331; Conservative 212; Mismatches 494; Indels 188; Gaps 43;

QY 1 ENLDTENAMTFQ-ENARGFQSVVOL---QGSRVVVVGAQEIVAAQNRGSLVQC--DYST 54
Db 30 YNGLPEAKIFSGSSEQYGAQQVQFINKGNWLLVGSFWSGPPENRMGDVYKCPVDLST 89
QY 55 GSCEPIRLQ-----VPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSENTYVKGLC 107

```

```

Db 90 ATCEKLNLOSTSIPIVNTKMTNMSLGLILTRNMGTGGELTCGLPLWAQCCNQYITGVC 149
QY 108 FIFGSLNRQOPQKPPALRGCPQEDSDIAFLIDGSGIIPHDPFRMKFVSTVMBQLK-- 165
Db 150 SDISPDF-QLSASFSPATQPCPSL-IDVVVVCDSESNISYPWD--AVKNFLEKFEVQGLDIG 205
QY 166 KSKTLFSLMOYSEEPFRIHETFEFQNNPNRSLVKPITQLLG-RTHATATVRKRVIRELN 224
Db 206 PKTQVGLQYANPRVFNLTNYKTEEMIVATSTQSYGDLTNTFGAIQYARKYAYS 265
QY 225 ITNGARKNAFKILIVITDGEKFGDPLGVEDYDIPEADREGVIRVIGV-----GDAPRESEK 279
Db 266 AASGGRSATAKVVVVVTDGESH-DGSMKXAVIDOCNHDNLRGLIAVLGVLNRLDITKN 324
QY 280 SRQELNTIASKPPRDHVQVNNFPEAKTTONOLREKIFAIEGTQTGSSSSFEHEMSQEGF 339
Db 325 LIKEIKAIASIPTEYFFNVNDEAALLERKAGTLGEQIFSTEGTVQG-GDNFQMEMSQVGF 383
QY 340 SAAITSNRP--LLSTVGSYDMAGVFLYTSKEKSTFINMT--RVSDSMN-DAILGYAAAI 394
Db 384 SADYSONDILMLGAVGAFGWSGTIVOKTSHGHLIFPKQAFQILODRNHSYLYGSVAA 443
QY 395 ILNRNVQSLVLAGAPRYQHIGLVAMFRQNTGMWESNANV-----KGTQIGAYFGASLCV 448
Db 444 ISTGESTHFVAGAPRANYTGQIVLYSVN-----ENGNIITVIOAHRGDQIGSYFGSVLCV 498
QY 449 DVDSNGSTDLVLIGAPHYIEQTR--GGQVSVCPPLRGQARWQCDVAVLGEQQPMGRFG 506
Db 499 DVDKOTITDVLVAGAPMNSDLKKEGRVYLFETIKKGILQHQ--FLEGPEGIENTRFG 555
QY 507 AALTVLGDVNGKLTDAITGAPCEDNRGAVLFGHTSGSGISPSHSORLAGS--KLSPR 564
Db 556 SAIAALSDILMDGFNDVIVGSPLENQNSGAVIYNGHQST-IRTKYSQILGSDGAFRSH 614
QY 565 LQYFGOSLGGQDLTMDGLVLDITGVAQGHVLLLRSQPLVRVRAIMEFNPREVARNVFECH 624
Db 615 LQYFGRLSDYGLDNGDSITDVISGAFQVQVQLWSQSIADVAIEASFTEKI--TLVNKN 672
QY 625 DQVVKGEAGEVVRVCLVHVKSTRDLREGGIISQVVTYDLALD---SGRPHSAFVNEFK 680
Db 673 AQII-----LKLCF---SAKFRPTKQNNQVAIVYNTILDADGFSRVTSGFLFKENN 721
QY 681 NSTRQTVLGTOTC--ETLKLQLPNCIEDPSPVILRLNFSLVGTPLSAFGLRPLVLA 738
Db 722 ERCLQKQMVVNAQSCPEHIITQEPS---DVVNSLDLRVDISLENPGTS-----PALE 772
QY 739 EDAQRLFTALFPPEKNCNDNITCQDDLSITF---SFMSLDCLVVGPPREFNVTVTRND 794
Db 773 AYSETAKVFSIPFKDCGEGDGLCISDLVLDVRQIPAAQEQPFIVSNQNKRLTFSVTLKN 832
QY 795 GEDSVTRQVTFPPDLDSYRKVSTLQNRQSRQSWRLACESAST-EVSGALKSTSCSINH 853
Db 833 RESAYNTGIVDFSENLEF-----ASFSLPVDGTETVCQVAASQKSVACDVGY 880
QY 854 PIFPENSEVTNTTFDVDSKASLGNKLLKANVTSENNMPTNKTEFOLELPVKYAVMV 913
Db 881 PALKREQQVTFINFDNLQ-NLQNASLSFOALSEQENKADNLVNLKIPLLYDAEI- 938
QY 914 VTSHGVSATKYNFTASENTSRVMQHQYQVSNLQOR-----SLPISLFLV----- 958
Db 939 ---HLTRSTNINFIYSSDGNVPSIVHSFEDVGPKFIFSLKVTGTSVPVSMATVHIHQ 995
QY 959 -----PVRLNQTVIWRPQVTF-SENLSSTCHTKER 988
Db 996 YTKENPLMYLTGVQTDKAGDISCNADINPLKIGT-----SSSVSPKSENFR---HTKE- 1047
QY 989 LPSHSDFLAELRKAPVNVCSIAVCORIQCDIPFFGIGQBEFNATLKNLSFDWIKTSHNH 1048
Db 1048 -----LNCRTASCNSVTCLWKDVKHMKGEYFVNVTVTRINWGTFFASTFTOT 1091
QY 1049 LLIVSTAETI-LFNDVSFTLLPGGAFVRSQETTKVPEFVNP-----LP--LIVGSSVG 1100

```





```
QY 645 STRDRREGQIQSVTVYDLALDSGR-----PHSRAVENETKNSRRROTQVLGLTQTCETL 699
Db 507 CFSYKGEVGYIVLFTNMSLDVNRKASPRPFYSSNGTSDVTGIGQSSREANCRT 566
QY 700 KLOLPNCIEDPVSPIVLRNLSFSLVGTPLS-----AFGNLRPVLAEADQR-LFTALFFPEK 753
Db 567 QAFMRKDVDRILTLPIQIEAAHYLGHVPHVISKRSTEEFPPLQPILOQKKEKIDMKKTINFAR 626
QY 754 NCGNDNICQDDLSIT--PSEM-----SLDCLVGGPREFNVTVTVNRNDEGDSYRQTQVTF 807
Db 627 FCAHEN--CSADLQWSAKIGLPHENKTYLAVGSMTKLMLNLSFNGAGDDAYETTLHVKL 685
QY 808 PLDLSYRKVSTLQNRQSRWRLACESASSTEVSGALKSTSCSNHPIFPENSVTNIT 867
Db 686 PVGLYFKILELEK-----QINCE---VTDNSGVVQ-LDCSICYIYVDHLSRIDISFL 735
QY 968 FVDVSKASLGNKLLKANVTSEN--NMPRTNKTQFQLELPVKYAVYVMVTSHGVS TKYL 925
Db 736 LDVSSLGRABEDLSITVHATCENEEDNKLKHSRVTVAIPLKYEVKLTVHGFVNPTSFVY 795
QY 926 FTASNTSRV-----MOHQVQSNLQORSIP-LSLVFLVPVRLNQTWIDRPPQVTFSEN 979
Db 796 GSDNENPETCWBEKXNLTHVINTGNSMAPNVSVEIMVNSFS-----PQTDKLFNI 848
QY 980 -----SSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDDIPFGIOBEFNATLK 1033
Db 849 LDVQTTTGBCHFEVQ-----RVCALEQ-----QKSAMQTLK 880
QY 1034 GNLSFDMYIKTSHNHLIVSTAE-----ILFN-----DSYFTLLPGGAFVRSQTE 1079
Db 881 GIVREF--LSKTRKLLAYCIKADPHCLNFCNFCMBSGKEASVHIQLEGRPS-ILEWDE 936
QY 1080 TKVPEFV-----PNP-----LPLVIGSSVGLL 1103
Db 937 TSALKFIRATGPEPNPRVIELNKDENVAHVLEGLHGHORPKRYFTVITVISSLLGLI 996
QY 1104 LLALITAAVLKGLFFKQYKDMWSE 1128
Db 997 VLLLSIYVMWKAGFFKQYKSIQOE 1021

RESULT 14
I58409
integrin alpha-9 chain precursor - human
C;Species: Homo sapiens (man)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C;Accession: I58409; A49459
R;Hibi, K.; Yamakawa, K.; Ueda, R.; Horio, Y.; Murata, Y.; Tamari, M.; Uchida, K.; Takah
Oncogene 9, 611-619, 1994
A;Title: Aberrant upregulation of a novel integrin alpha subunit gene at 3p21.3 in small
A;Reference number: I58409; MUID:94119603; PMID:8290272
A;Accession: I58409
A;Status: preliminary; translated from GE/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1035 <RES>
A;Cross-references: UNIPROT:Q13797; GB:D25303; NID:g464180; PIDN:BARA04984.1; PID:g533327
R;Palmer, E.L.; Ruegg, C.; Ferrando, R.; Pytela, R.; Sheppard, D.
J. Cell Biol. 123, 1289-1297, 1993
A;Title: Sequence and tissue distribution of the integrin alpha9 subunit, a novel partne
A;Reference number: A49459; MUID:94064789; PMID:8245132
A;Accession: A49459
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 30-1035 <PAL>
A;Cross-references: GB:I24158
C;Superfamily: integrin alpha-4 chain
C;Keywords: glycoprotein; metal binding; transmembrane protein
F;1-27/Domain: signal sequence #status predicted <SIG>

Query Match 10.7%; Score 630; DB 2; Length 1035;
Best Local Similarity 26.5%; Pred. No. 7.5e-35;
Matches 237; Conservative 160; Mismatches 332; Indels 164; Gaps 40;
```

```
QY 331 EHEMSOEGFSAATISNGPLLSVTGSDWAGGVFLYTSKEKSTFINMT-RVDSMDNDAYLG 389
Db 190 EHSCQAGIAGFTTEELVVMGAPGSGFYWAGTIKVNLITD-NYKLKNDENVIMNRYTYLG 248
QY 390 YA--AAIILNRNVQSLVLGAPRYOHIGLVAMFR--QNTGMWESNANVKGQIQAIFASL 445
Db 249 YAVTAGHFSHPSTIDVVGGAPQDKIGKVYIFRADRRSGTLIKIFQASGKMGSYFGSSL 308
QY 446 CSVDVDSNGSTDVLIGAPHYEQTRGGQVSVCPLPRGQORWQCDVAVLGEOQPWGRF 505
Db 309 CAVDLNGDGLSDDL-LVGAPMFEIRDEGQVTYV--INRNGGALAE-QUALTGD-GAYNAHF 364
QY 506 GAALTVLGDNVNGDKLTDVAICAGEEDNRGAVLFGHTSGSGISPSHSORIASGKLSPL 565
Db 365 GESIASLDDLDNDGFPDVAIGAKEDDFAGAVIYHGDAG-GIVPQYSMLKSGKINPVL 423
QY 566 QYFQSGLSGGODLTMDGLDVTGA--QGHVLLRSOPVLVRVKAIMEFNPREVARNVPEC 623
Db 424 RMFQSGISGIGDMDGNGYDPDVTVGAFMSDSVLLRARPVITVD-VSIFLPGSINITAPQC 482
QY 624 NDQVVKGEAGEVRV-CLHVQ-----KSTRDLREGQIQSVTVYDLALDSGRPHRAV 675
Db 483 HD-----GQOPVNCVLTTCFSFHGHVPBEI---GLNYVLMAVAKKEKQMRPVY 531
QY 676 F--NETKNSTRROTQVLGLTQTCETLKLQLPNCIEDPVSPIVLRNLSFSL---VGTPLS 728
Db 532 FVLGETMGQVTEKLQLTMYEETCRHYVAHVKKRVQDVSPIVEAAYSEHVTEGEER 591
QY 729 AFGNLRPVL-----AEDAQRLFTALPFPEKNCNDNITCQDDLSITFSFM--SLD----CL 777
Db 592 ELPLPTEVLRWKKGOKIAQKNQTV---FERNCRSED-CAADLQLQGLKLLSSMDKTYL 647
QY 778 VVGPREFNVTVTVNRNDEGDSYRQTQVTFPFLDLSYRKVSTLQNRQSRWRLACESASS 837
Db 648 ALGAVKNISLNISISNLGDDAYDANVSFNVSRELFF-----INMOKEMGISCELLS 701
QY 838 TEVSGALKSTSCSNHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPRTNK 897
Db 702 DFL-----KCSVGFPPMRSKSYEFVIFDTSHL--GEEVLSFVITVQSG--NTER 750
QY 898 TE-----FOLELPVKYAVYVMVTS-----HGVSTKYNFTASENTS---RVMOHQY 940
Db 751 SESLHDNTLVLMVPLMHEVDTSITGIMSPTSFSVYGESVDAANFTQLDDLECHFPINITL 810
QY 941 QVSNLQORSIPISLVFL-VPVRLN-----QTVIWDPRQVTVSENLSSTCHTKER 988
Db 811 QVYNTGPTSLPGSSVSISFNNRSLSSGGAEMFHVQEMVVGQEKNCFSQKNPTPCIIPQEQ 870
QY 989 LPSHSDFLAELRKA-----PVCNCSIAVCORIQCDDIPFGIOBEFNATLKNLSFD 1039
Db 871 ENIFHTIFAFKTSGRKVLDCXKPGISCLTAHCN-----FSALAKEESRTI-----D 917
QY 1040 WYIKTSHNHLIVSTAEILFNDSVFTLLPGGAFVRSQTEKTEVP----- 1084
Db 918 IY-----MLLNT-BILKDDSSVIQ-----FMSRAKVKVDPALRVVETIAHGNPEEV 962
QY 1085 ---FEVFN---PPLPIVG-----SSVCGLLLLALITAAVLKGLFFKQYKDMW 1126
Db 963 TVVFEALHNLEPRGYVVGWIIATSLVGLIDIFLLAVLLMKMGFFRRRYKEII 1015

RESULT 15
T31437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C;Species: Lytechinus variegatus (variegated urchin)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C;Accession: T31437
R;Hertzler, P.L.; McClay, D.R.
Submitted to the EMBL Data Library, May 1998
A;Description: Alpha SU2, a sea urchin integrin which binds laminin.
A;Reference number: 421035
A;Accession: T31437
```



A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1041 <HEP>

A:Cross-references: UNIPROT:O76378; EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC2

A:Experimental source: developmental stage embryo

C:Function:

A:Description: binds laminin

C:Superfamily: integrin alpha-2b chain

Query Match 10.5%; Score 614.5; DB 2; Length 1041;

Best Local Similarity 24.8%; Pred. No. 8.8e-34;

Matches 236; Conservative 155; Mismatches 314; Indels 245; Gaps 41;

```
QY 336 QRGFSAAITSNGP--LLSTVGSYDWDAGGVFLYTSKEKSTFINNTRVDS-----DMNDAYLG 389
Db 181 QAGFGIIFSDNSALVWAGPGSYLQGIYVQSLNRSV-VQATQESNTGTYSFDNSYRG 239

QY 390 YAAAI--ILNRVQSILVLAGAPYQHI-GLVAMPRQNTGMWESNANVKGTOIGAYFGASLC 446
Db 240 YSLALGDFNGGVQDYVVGTPRAESLWGLVAIFDQNLNQFN---QVMGTQIVAYFGYSVT 296

QY 447 SYDVDSNGSTDVLIGAPHYEQTRGGQSVSCPLPRGQRARMOCDAVLYGEQ-----498
Db 297 VVDI-NNDTYDOLLVGA PMYDDGPAIQ-----RWEAGAVVYVLQNPVGPGA 343

QY 499 -----GPGWRFGAALTVLGDVNGDKLTDVAIGAGEEDNRGAVYLFHGTSGS 546
Db 344 SNRLSUSLTIIGQIKRSRFLGSLTASIGDSNQDGFNDVAIGAPYEGDAGAVYIYHG-SAN 402

QY 547 GISPSHSQRIAGSKLS-PRLOQYFGQSLGGQDLTMDGLVDLTVGAQ--GHVLLRSQPVL 603
Db 403 GLKSTPAQVLTPSTLGHSGITTFGSLQGGQMDKKNYPDLLVGAESANTAVLIRTPVV 462

QY 604 RVKAIMFNPREVARNVFECDNVQVKGAGEVRVCLHVQKSTRDLREGQIQSV-----558
Db 463 SLDATLNTPE-----IGINLENKTYE-LADGTMTVTSFIAMT 497

QY 659 -----VTYDLALDSG-RPHSRVAFNETKNSRRQTVGLL-TQTCETLKLQ 702
Db 498 CFTYTGNYLPDHIDISYTVTVDSGIITANRRAMEVDMDMSEITKTRRLAVSTQFCDELRAY 557

QY 703 LPNCIEDPVSPVLRNLFSLVGTPLSAFGN-----LRPVLAEADAQRLETFALPFPFEK 753
Db 558 VGNSIEDKLTPIKVTIQLYDL-----NNDESRLQPHLELPIDMATMSTQTKQVSIQN 609

QY 754 NCGNDNICODDLSITFSFMSLCLVGGPREFNVTVVRNDGEDSYRTQVTFPFDLSY 813
Db 610 NCVN-NICIPDLDTVVT-PNLEIVIGTQELTLDVSLNRRGEDAFQSSLSVYVPLGLQF 667

QY 814 RKVSTLQNRQSRWRLACESASSTEVSGALKSTCSINHPIPPEN-----SEVTFNIT 867
Db 668 VRI-----ERRANMDFSVTCESD-----LRITCDTGNPMVGKNILBFLTLSTFQVS 717

QY 868 FDVDS-----KASLGNKLLKANVTSENMPRTNKTEFQLELPKYAVYVMVVTSHGVSTK 922
Db 718 GDKDSIEFYFKAESENS--EDPNTLENNELNMTVPVTVDCTLKLLSASYPEIWMYSTQED 775

QY 923 YL-----NFTASNTSRVMQHOYQV-----SNLQORSLPIS-----LVFLVPV 960
Db 776 YVVPFPFPAKNASPADIGMEVM-HLYEVRNTGSSNAGEVSLNIQWPKNEDGEYLFYLLGI 834

QY 961 RUNQTVIMDRPO-----VTFSENLSTCHTKELPSHSDFLAELRKAPVV 1005
Db 835 MTEEGVTCQLTQKAMPEGVKLPSTKAKLSNSTTVQSGRKRREPEVAEALAQDN--VI 892

QY 1006 NCSIAVCORIQCDDIPFFGQEEFNAT-----LKGNLSFDWYIKTSHNHLLIVSTAELIF 1059
Db 893 YCASDSCVLINCTI-----DEINASKSVVIRILGRF---W-----ERTF 928

QY 1060 NDSVFTLLPQGGAFVRSQIETKVE--PFEVPNP-----LP-----1092
Db 929 QKAVSELTPVQVQATIASASAAVKTIPIYNIPLPRDFSDSKASTLVTBELVPPVPIAW 988
```

QY 1093 -LIVGSSVGGLLLLLALITAAKYKLGFFKR-----QYKDMMS--EGGPP 1132

Db 989 WIIIVSVLGGIILLIILILGLWACGFFERKKPCEKEKYPAPVASADKDGPP 1038

Search completed: November 9, 2004, 12:19:40

Job time: 33.25 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:46:42 ; Search time 136.75 Seconds  
(without alignments)  
4783.919 Million cell updates/sec

Title: US-09-902-481b-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5852.5	99.6	1152	1 ITAM_HUMAN	P11215 homo sapien
2	4476	76.2	1153	1 ITAM_MOUSE	P05555 mus musculus
3	4387	74.7	1151	2 Q9J130	Q9J130 rattus norv
4	3815.5	64.9	1920	2 Q28984	Q28984 sus scrofa
5	3483	59.3	1163	1 ITAX_HUMAN	P20702 homo sapien
6	3417	58.2	1162	1 ITAD_HUMAN	Q13349 homo sapien
7	3335.5	56.8	1188	2 Q6KAS4	Q6KAS4 mus musculus
8	3335.5	56.8	1188	2 BAD21383	Bad21383 mus muscu
9	3326.5	56.6	1169	1 ITAX_MOUSE	Q9QXH4 mus musculus
10	3251.5	55.3	1161	1 ITAD_RAT	Q9QYE7 rattus norv
11	1578	26.9	1165	1 ITAL_BOVIN	P61625 bos taurus
12	1552.5	26.4	1170	1 ITAL_HUMAN	P20701 homo sapien
13	1552	26.4	1166	2 Q6TVB8	Q6TVB8 bos taurus
14	1552	26.4	1166	2 AAQ90015	Aaq90015 bos tauru
15	1533.5	26.1	1163	1 ITAL_MOUSE	P24063 mus musculus
16	1529.5	26.0	1161	2 Q9WTV4	Q9WTV4 mus musculus
17	1519	25.9	1160	2 Q9R200	Q9R200 mus musculus
18	1414	24.1	1196	2 Q9RTF1	Q9RTF1 cyprinus ca
19	1363.5	23.2	1187	2 Q9RTF0	Q98tf0 cyprinus ca
20	1350.5	23.0	1086	2 Q96HB1	Q96hb1 homo sapien
21	1270	21.6	927	2 Q8HZV0	Q8hzv0 bos taurus
22	1163.5	19.8	1167	2 Q8G340	Q8G340 rattus norv
23	1157.5	19.7	1167	1 ITAE_MOUSE	Q88340 rattus norv
24	1149	19.6	1179	1 ITAE_HUMAN	P38570 homo sapien
25	1119	19.0	1167	2 Q88341	Q88341 rattus norv
26	1102.5	18.8	1151	1 ITA1_HUMAN	P56199 homo sapien
27	1093.5	18.6	1189	1 ITAH_HUMAN	Q9UXK5 homo sapien
28	1084	18.5	1170	1 ITAZ_BOVIN	P53710 bos taurus
29	1072	18.2	1178	1 ITA2_MOUSE	Q62469 mus musculus
30	1072	18.2	1178	2 Q6P1C7	Q6pic7 mus musculus
31	1072	18.2	1178	2 AAH65139	Aah65139 mus muscu

#### ALIGNMENTS

##### RESULT 1

ITAM_HUMAN	ITAM_HUMAN	STANDARD;	PRT;	1152 AA.
ID	P11215;			
DT	01-JUL-1989 (Rel. 11, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha subunit) (CR-3 alpha chain) (CD11b) (leukocyte adhesion receptor MO1)			
DE	(Neutrophil adherence receptor)			
GN	Names:ITGAM; Synonyms:CR3A, CD11B;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID:9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88315033; PubMed=2457584;			
RA	Corbi A.L., Kishimoto T.K., Miller L.J., Springer T.A.;			
RT	"The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b) alpha subunit. Cloning, primary structure, and relation to the integrins, von Willebrand factor and factor B.";			
RL	J. Biol. Chem. 263:12403-12411(1988).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88190151; PubMed=2833753;			
RA	Arnaut M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;			
RT	"Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion glycoprotein Mol: chromosomal localization and homology to the alpha subunits of integrins.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=88257215; PubMed=2454931;			
RA	Arnaut M.A., Gupta S.K., Pierce M.W., Tenen D.G.;			
RT	"Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mol (complement receptor type 3).";			
RL	J. Cell Biol. 106:2153-2158(1988).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=93123748; PubMed=8419480;			
RA	Fleming J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;			
RT	"Structural analysis of the CD11b gene and phylogenetic analysis of the alpha-integrin gene family demonstrate remarkable conservation of genomic organization and suggest early diversification during evolution.";			
RL	J. Immunol. 150:480-490(1993).			
RN	[5]			
RP	SEQUENCE OF 9-1153 FROM N.A.			
RX	MEDLINE=89098993; PubMed=2563162;			
RA	Hickstein D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L., Roth G.J.;			
RT	"cDNA sequence for the alpha M subunit of the human neutrophil adherence receptor indicates homology to integrin alpha subunits.";			

P17301 homo sapien  
O75578 homo sapien  
P18614 rattus norv  
P61622 mus musculus  
Q77QC3 mus musculus  
O42094 gallus galli  
Q8BS01 mus musculus  
Q9WU8 mus sp. itg  
Q8MKF4 felis silve  
Q8TES5 homo sapien  
Q9BPQ8 halocynthia  
Q63001 rattus norv  
Q6P666 mus musculus  
AAH57200 mus muscu

32 1069 18.2 1181 1 ITA2\_HUMAN  
33 1065.5 18.1 1167 1 ITAG\_HUMAN  
34 1060 18.0 1180 1 ITAI\_RAT  
35 1052.5 17.9 1188 1 ITAH\_MOUSE  
36 1052.5 17.9 1188 2 Q77QC3  
37 1051 17.9 1171 2 O42094  
38 1041 17.7 1038 2 Q8BS01  
39 1014 17.3 895 2 Q9WU8  
40 1013.5 17.3 1160 2 Q8MKF4  
41 865 14.7 348 2 Q8TES5  
42 852 14.5 1332 2 Q9BPQ8  
43 807 13.7 205 2 Q63001  
44 784.5 13.4 304 2 Q6P666  
45 784.5 13.4 304 2 AAH57200





RT complement receptor type 3 and cellular adhesion molecule Mac-1. ";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5644-5648 (1986).  
 RP [4]  
 RN SEQUENCE OF 17-28.  
 RX MEDLINE=85188276; PubMed=3887182;  
 RA Springer T.A., Teplov D.B., Dreyer W.J.;  
 RT "Sequence homology of the LFA-1 and Mac-1 leukocyte adhesion  
 glycoproteins and unexpected relation to leukocyte interferon.";  
 RL Nature 314:540-542 (1985).  
 CC -!- FUNCTION: Integrin alpha-M/beta-2 is implicated in various  
 adhesive interactions of monocytes, macrophages and granulocytes  
 as well as in mediating the uptake of complement-coated particles.  
 CC It is identical with CR-3, the receptor for the iC3b fragment of  
 CC the third complement component. It probably recognizes the R-G-D  
 CC peptide in C3b. Integrin alpha-M/beta-2 is also a receptor for  
 CC fibrinogen, factor X and ICAM1. It recognizes P1 and P2 peptides  
 CC of fibrinogen gamma chain. Alpha-M/beta-2 play a critical role in  
 CC mast cell development and in immune complex-mediated  
 CC glomerulonephritis. Mice expressing a null mutation of the alpha-M  
 CC subunit gene demonstrate increase in neutrophil accumulation, in  
 CC response to a impaired degranulation and phagocytosis, events that  
 CC apparently accelerate apoptosis in neutrophils. These mice develop  
 CC obesity.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta chain. Alpha-M  
 CC associates with beta-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=1;  
 CC IsoId=P05555-1; Sequence=Displayed;  
 CC Name=2;  
 CC IsoId=P05555-2; Sequence=VSP\_010473;  
 CC Note=No experimental confirmation available;  
 CC -!- TISSUE SPECIFICITY: Predominantly expressed in monocytes and  
 CC granulocytes.  
 CC -!- DOMAIN: The integrin I-domain (inset) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; X07640; CAA30479.1; -;  
 DR EMBL; AK039444; BAC30350.1; -;  
 DR EMBL; M14293; AAA39484.1; -;  
 DR PIR; S00551; S00551.  
 DR HSSP; P11215; 1BHQ.  
 DR MGP; MGI:96607; Itgam.  
 DR GO; GO:0009897; C:external side of plasma membrane; IDA.  
 DR GO; GO:0007155; P:cell adhesion; IMP.  
 DR GO; GO:0045123; P:cellular extravasation; IMP.  
 DR GO; GO:0030593; P:neutrophil chemotaxis; IMP.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF01839; FG-GAP; 3-  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWFA; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 KW Alternative splicing; Calcium; Cell adhesion;  
 KW Direct protein sequencing; Glycoprotein; Integrin; Receptor; Repeat;  
 KW Signal; Transmembrane.

FT	SIGNAL	1	16	
FT	CHAIN	17	1153	Integrin alpha-M.
FT	DOMAIN	17	1105	Extracellular (Potential).
FT	TRANSMEM	1106	1129	Potential.
FT	DOMAIN	1130	1153	Cytoplasmic (Potential).
FT	REPEAT	31	84	FG-GAP 1.
FT	REPEAT	85	163	FG-GAP 2.
FT	DOMAIN	164	350	VWFA.
FT	REPEAT	337	400	FG-GAP 3.
FT	REPEAT	401	452	FG-GAP 4.
FT	REPEAT	454	515	FG-GAP 5.
FT	REPEAT	517	575	FG-GAP 6.
FT	REPEAT	580	632	FG-GAP 7.
FT	CA_BIND	465	473	Potential.
FT	CA_BIND	529	537	Potential.
FT	CA_BIND	592	600	Potential.
FT	SITE	1132	1136	GFFKR motif.
FT	DISULFID	66	73	By similarity.
FT	DISULFID	105	123	By similarity.
FT	DISULFID	654	711	By similarity.
FT	DISULFID	770	776	By similarity.
FT	DISULFID	999	1023	By similarity.
FT	DISULFID	1028	1033	By similarity.
FT	CARBOHYD	58	58	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	86	86	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	391	391	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	696	696	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	734	734	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	772	772	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	801	801	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	881	881	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	907	907	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	941	941	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	980	980	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	994	994	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	1022	1022	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	1045	1045	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	1051	1051	N-linked (GLCNAC. .) (Potential).
FT	CARBOHYD	1076	1076	N-linked (GLCNAC. .) (Potential).
FT	VARSPIC	453	569	Missing (in isoform 2). /FTID=VSP_010473.
FT	CONFLICT	37	37	N -> S (in Ref. 2).
FT	CONFLICT	683	683	V -> G (in Ref. 2).
FT	SEQUENCE	1153	AA; 127480	MW; 178DB988AECB0343 CRC64;

Query Match 76.28; Score 4476; DB 1; Length 1153;  
 Best Local Similarity 74.3; Pred. No. 2.4e-287;  
 Matches 845; Conservative 142; Mismatches 149; Indels 2; Gaps 2;

Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVANQSGSLYQCDYSTGSCBPI	60
Qy	17	FNLDTENAMTFQENARGFGQSVVQLQGSRVVVGAPQEIIVANQSGSLYQCDYSTGSCBPI	76
Qy	61	RLQVPVAVNMSLGLSLAATTPSPQLLACQPTVHQTCSNTYVKGCLFGLFSGNLRQOPQK	120
Qy	77	PLQVPVAVNMSLGLSLAATTPSPQLLACQPTVHQTCSNTYVKGCLFGLFSGNLRQOPQK	136
Qy	121	FPEALRCPOEDSDIAFLIDGSGIIIPDPRMKEFVSTWMEQLKSKTFLPSLMQYSEEF	180
Qy	137	FPEALRCPOEDSDIAFLIDGSGIIIPDPRMKEFVSTWMEQLKSKTFLPSLMQYSEEF	196
Qy	181	RIHFTFKFQNNPRLSVKPIITOLLGRTHATCVRVIRELLNITNGARKNAFKILIVI	240
Qy	197	RIHFTFKFQNNPRLSVKPIITOLLGRTHATCVRVIRELLNITNGARKNAFKILIVI	256
Qy	241	TGGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQELUNTATSKPRDHVFOVN	300
Qy	257	TGGEKFGDPLGYEDVIEADREGVIRVIVGVDAFRSEKSRQELUNTATSKPRDHVFOVN	316
Qy	301	NFEALKTIONLREKIFAIETGOTGSSSFEHEMSQEGSAATNSGNPLSTVGSYDWAG	360
Qy	317	NFEALKTIONLREKIFAIETGOTGSSSFEHEMSQEGSAATNSGNPLSTVGSYDWAG	376





Qy	781	GPREFNVTVVRNDGDESYRTQVTFPPDLSTRYKTVSTLONQSRQSRWRLACESASSTEV	840
Db	796	DSRDFDSVTLVRNDGDESYTKVTCYPSGLSYRKYVSASQNFSSKXPRVIAE-PSSEGG	854
Qy	841	SGALKSTCSINHPFIPENSEVFNTFTFDVDSKASLGNKLLLKANTVSENMMPTNKTEF	900
Db	855	QGVLKSTIWDINHPIIPANSEVFNTFTFDVDSASLGNKLLLVVVTSENNVSRTDKTES	914
Qy	901	QLELPVUYAYVMVTVSHGVSTKYLNFTASENTRGVMOHQVSNLQORSPLPSLVFLVPV	960
Db	915	QLELPVUYAYVMVTVSESSIKYLNFTASMTSKVQHQYQFNLQORSFVSVFWFVPIV	974
Qy	961	RLNQTVIWDPPQVTFSSNLSTCHTKERLPSHDSFLAELRKAPVWVNCIAVQCRIQCDIP	1020
Db	975	QINKVTIWDPPQVTFSQLSSVCRTEQKSPSHKQFDELRTPVLNCSVAVCKRIQCDIP	1034
Qy	1021	FRGIOEENFATLKNLSFDWKYIKTSHNHLLIVSTAELLFNDSVFTLLPQOCQAVRQOTET	1080
Db	1035	SFNSKEIFNVTLQGNLLFDWYIETSHDHLVLVSTAELLFNDSAFALLPGQETFVKQACT	1094
Qy	1081	KVPFFEPNPLPIVGVSSVGLLLALITAAALYKLGFFKRYQKDMMSBEGPPGAEPQ	1137
Db	1095	KVPFPYVTHNVPVPIVGVSSVGLVLLAITAGLYKLGFFKRYQKDMMAEAGODGPPQ	1151

Qy	238	IVITDEKFGDPLGYEDVPIEADREGVIRYVIGVDAPSEKSRQBELNLTIA SKPPDPHFV	297
Db	121	VITDEKFGDPLGYEDVPIEADRKGVIRYVIGVDAPFNSWKREELNLTIA SKPCGDHFV	180
Qy	298	QVNNFALKTIQNOLREKIFAETGTGTGSSSSPEHEMSQEGFSAAITSGNPLLS	357
Db	181	QVNNFRAVKTIQNOLQEKIFAETGTGTSTSPCEMSQEGFSAAITSGNPLLS	240
Qy	358	WAGGVFLYTSKEKSTFINMTRVDSMNDAYLGAVALILRNRVQSLVILGAPRYQHIL	417
Db	241	WAGGAFLEHMPKRVIFINTTRVDSMNDAYLGAVALILRNQVSLVILGAPRYQHTGL	300
Qy	418	MFQNTGMNESNANVGTQIGAYFGASLCSVDVDSNGSTDVLILGAPHYYEOTRGQ	477
Db	301	MFQNSGAMKEKNADIKGSQIGSYFGASLCSVDVNRGSDVLILGAPHYYEOTRG	360
Qy	478	CPLPRGORARWQCDVILYGEQGPWRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR	537
Db	361	CPLPQG-RAKWCRCRVILCGEQHPWRSRFGAALTALGDVNGDKLTDVAIGAPGE	419
Qy	538	YLPHGTSGSGISPSHSQRTAGSKLSRLOYFGOSLSGGQDLTMDGLVDTVGAQGH	597
Db	420	YLPHGTSELGISPSHSQRTAGSQLSRLOYFGOSLSGGQDLTMDGLMDLAVGAQGH	479
Qy	598	RSPQVLRVKAIMEFNPREVARNVFCGNDQVVGKEAGEVRVCLHVOKSTRDLREGQ	657
Db	480	RSPQVLRVKAIMEFNPREVARNVYECQQAATQIAGEVQVCLVOKSTWDLREGDTQ	539
Qy	658	VVTYDLALDSGRPHSRVAVNETKNSRTROTVLGLTQTCETLKLQLPNCIEDPVS	717
Db	540	IITYDLALDPRPHPRVAFEBETKNTRTOTLGLSRKCEHLALWLPDCVEDSVTP	599
Qy	718	LNFSLVGCTPLSAPGNLRPVLAEDAQRLFTALPPEKNCMGNDICQDDLSITPFS	777
Db	600	LNFSLVGKPASSGNLRPVLAQAQLFTALPPEKNCMGNDICQDDLSITPFSMSL	659
Qy	778	VVGQPREFNVTVRNMGDSYRTQVTFPPDLDSYRKVYSTLQNSQRSWRLACPSAS	837
Db	660	VVGQPRDLKVTLLVRNQGEDSYRTQVTFPPDLSYRKVYSTQNSQRSWRLACPSD	719
Qy	838	TEVSGALKSTSCINHPPIPENSEVFNITEDVDSKASLGNKLLKANVTSENMPRT	897
Db	720	TEBSTALKSTSCINHPPIPDNSEVFNTVDVNPDAFLGYKULLKANVTSENMPSS	779
Qy	898	TEFQLELPVKYAYVMVVTSHGVSTKYLNFNTASENTSRVMQHOYQVSNLQORS	957
Db	780	TEFQLELPVKYAYVVVVTSLVSTKYFNFTASEKTRHVIEHQYQFNGLGQRKL	839
Qy	958	VPVRLNQTVIWDPRQVTFSENLSSCTHKERLPSHSDFLAHLKAPVWVNCISIA	1017
Db	840	VPVRLNRTVMDQVQVTFSQLSRGSCSTEBIGRHSDFLEKLOKTPVLNCSIAVC	899
Qy	1018	DIPFPGIOEBFNATLKNLSF 1038	
Db	900	DIPFSGIOELKVTILKNLSF 920	
RESULT 5			
ITAX_HUMAN			
ID	ITAX_HUMAN	STANDARD;	PRT; 1163 AA.
AC	P20702; O8IVA6;		
DT	01-FEB-1991 (Rel. 17, Created)		
DT	05-JUL-2004 (Rel. 44, Last sequence update)		
DT	01-OCT-2004 (Rel. 45, Last annotation update)		
DE	Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95		
DE	alpha chain) (leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).		
GN	Name=ITGAX; Synonyms=CD11c;		
OS	Homo sapiens (Human)		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_Taxid=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		



```
FT TURN 160 161
FT HELIX 164 178
FT TURN 179 180
FT TURN 183 185
FT STRAND 186 193
FT STRAND 197 201
FT HELIX 203 208
FT HELIX 212 216
FT TURN 217 218
FT STRAND 226 226
FT HELIX 228 236
FT TURN 237 240
FT HELIX 242 244
FT TURN 245 245
FT TURN 248 249
FT STRAND 251 258
FT STRAND 263 263
FT HELIX 269 278
FT TURN 279 280
FT STRAND 282 288
FT HELIX 290 293
FT TURN 296 297
FT HELIX 298 304
FT HELIX 310 312
FT STRAND 313 316
FT HELIX 319 325
FT HELIX 326 334
FT TURN 335 335
SQ SEQUENCE 1163 AA; 127886 MW; 83658A13B5C5DE8F CRC64;

Query Match
Best Local Similarity 59.3%; Score 3483; DB 1; Length 1163;
Matches 693; Conservative 138; Mismatches 292; Indels 6; Gaps 4;

Qy 1 PNLDTENAMTFOENARGQGVSVQLQGSVVVVGAPQEIIVANQSGSYQCDYSTGSCPEI 60
Db 20 PNLDTTELTFRVDSAGFGDSVGVYANSVVVVGAPQKITAAQOTGGLYCGYSTGACEPI 79

Qy 61 RLOVPVAVNMSLGLSLAATTSPPQLLACGPTVHTQTSENTYVKGLCFLFSNLRQPOK 120
Db 80 GLQVPPVAVNMSLGLSLAATTSPPQLLACGPTVHHECGRNMYLTGLCLLGP--QLTOR 137

Qy 121 FPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKEFVSTVMEQLKSKTKLFLSLMQYSEEF 180
Db 138 LPVSRQCPREQDIFVLIDGSGSISSRNFAFMNFVRAVISQRPSTQPSLQFQSNKF 197

Qy 181 RIHTTFEQNNPNRSILVKPITQLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
Db 198 QTHFTFEFRFTSNPLSLLASVHQLQGFTYATATQNVVHRLFHASYGARRDATKILIVI 257

Qy 241 TDGKEFGDPLGYEDVIPADREGVIRVYGVGDAPRSEKSPQELNTTASKPPRDHVFQVN 300
Db 258 TDGKKGDSLDYKDVIPWADAAGIIRVAIGVGLAFQNRNSWKELNDTASKPSQEHIFKVE 317

Qy 301 NFEALKTIQNLREKIFAIEGTOTGSSSSFFHEMSQGFSAATISNGPILLSVGSYDWAG 360
Db 318 DFDALKDIONLQKEKIFAIEGTETSSSFELEMAQEGFSAVFTPDGPVLGAVGFSFTWSG 377

Qy 361 GVFLYTSKEKSTFNTMRVDSMDMDAYLGAAAAIILNRVQSVLGVAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFINMSQENVDMRDSVLGYSTELALWKGVQSLVLGAPRYOHTKAVIFT 437

Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEQTRGGGVSCPL 480
Db 438 QVSRQWRKAEVGTQIGSYFGASLCSVDVDTGSTDVLIGAPHYEQTRGGGVSCVCLP 497

Qy 481 PRGORARQCDVAVLYGEGQGPWGRFGAALTVLGVNVDGDKLTDVAIGAPEGEDNRGAYLYF 540
Db 498 PRGWR-RWVCDVAVLYGEGQGPWGRFGAALTVLGVNVDGDKLTDVVIAGPEENRGAYLYF 556

Qy 541 HGTSGSGISPSHSORIASGSKLSPLQYFGQSLSGQDLTMDGLVDLTLVGAQCHVILLRSQ 600
Db 557 HGVLPSPISPSHSORIASGSKLSRLQYFGQALSGGQDLTQDGLVDLAVGARGQVLLLRTR 616
```

```
Qy 601 PVLEVKAIMBEFNPREVARNVPECNDOYVVKGEAGEVVRVCLHVQKSTRDRREGIOISVVT 660
Db 617 PVLWVGVMQFIPAEIIPRSAPFECREQVVSQETLVQSNICLYIDKRSNLCGRDLQSSVT 676

Qy 661 YDLALDSGRPHSRAVFNETKSNSTRQTVLGLTCTCTLKLQLPNCIEDPVSPVLRINF 720
Db 677 LDALDPGRLLSPRATFOETKNRSLSRVRLGLKAHCENFNLLLPSCVEDSVTPITLRNF 736

Qy 721 SLVGTPLSAFGNLAPVLAEDAQRLLFTALFPFEKNCNGDNICODDLSTIFFSMLSCLVVG 780
Db 737 TLVGKPLLAFRNLAPMLAADAQRYFTASLPFEKNCGADHICQDNLGHSFSFPGKSLVVG 796

Qy 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYKVKVSTLQNRQSRQSRWLACESASSTEV 840
Db 797 SNLENAEVMVWVNDGEDSYGTTITFSHPAGLSYRYVAGQKQGRSLHLTCDAPVG-- 854

Qy 841 SGALKSTSCSINHPIFFENSEVENTIITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900
Db 855 SQGTWSTSCRINHLIFRGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPTSKTTF 914

Qy 901 QLELPVKVAVYVVTSHGVSTKYLNFTAS-ENTSRVMQHQYOVSNLQORSIPISLVLVLP 959
Db 915 QLELPVKVAVYVTVSSHEQFTKYLNFSEBEKESHVAMHRYQVNNLQORDLPVSIWVWP 974

Qy 960 VRLNOTVIMDRPOVTFSENLSSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIOCDI 1019
Db 975 VELNQEAVMVDVESHQPNFSLRCSSEKIAPPASDFLAHQKVPVLOCSITAGLRFRCDV 1034

Qy 1020 PFGIOBEFNATLKGNSLFDWYIKTSHNHLIVSTAELFNDSTVFTLLPGQGAFFVRSQTE 1079
Db 1035 PFSFVQELDFTLLKGNLSPGWVRQILQKKVSVVSVABITPDSVYSQLPQGAFFMRAQT 1094

Qy 1080 TKVEPFPVNPPLPIVGVSSVGGLLLLLALITAAALYKLGFFFRQYKMMSE 1128
Db 1095 TVLEKYVHNPTPLIVGSSIGGLLLALITAVLYKVGFFFRQYKEMMEE 1143

RESULT 6
ITAD HUMAN
ID ITAD HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN Name=ITGAD;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=96111956; PubMed=8777714;
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT 3.";
RL Immunity 3:683-690 (1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin
RT gene CD11d. Essential role of Sps and Sp3.";
RN [3]
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., Lebeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human
RT beta 2-integrin alpha subunit.";
```

RL Gene 171.291-294 (1996).  
 RN [4]  
 RP INTERACTION WITH VCAM1.  
 RX MEDLINE:95059842; PubMed:9841932;  
 RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,  
 RA Hoffman P.A., Staunton D.E., Bochner B.S.;  
 RT "alphabeta2 integrin is expressed on human eosinophils and functions  
 as an alternative ligand for vascular cell adhesion molecule 1 (VCAM-  
 1)";  
 RL J. Exp. Med. 188:2187-2191 (1998).  
 RN [5]  
 RP INTERACTION WITH VCAM1.  
 RX MEDLINE:99370002; PubMed:10438935;  
 RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,  
 RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;  
 RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a  
 binding interface between I domain and VCAM-1";  
 RL J. Immunol. 163:1984-1990 (1999).  
 CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and  
 VCAM1. May play a role in the atherosclerotic process such as  
 clearing lipoproteins from plaques and in phagocytosis of blood-  
 borne pathogens, particulate matter, and senescent erythrocytes  
 from the blood.  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D  
 associates with beta-2.  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed moderately on myelomonocytic cell  
 lines and subsets of peripheral blood leukocytes and strongly on  
 tissue-specialized cells, including macrophages foam cells within  
 atherosclerotic plaques, and on splenic red pulp macrophages.  
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U37028; AAB38547.1; -;  
 DR EMBL: U40274; AAB60634.1; -;  
 DR EMBL: U40275; AAB60635.1; -;  
 DR EMBL: U40276; AAB60636.1; -;  
 DR EMBL: U40277; AAB60637.1; -;  
 DR EMBL: U40278; AAB60638.1; -;  
 DR EMBL: U40279; AAB60639.1; -;  
 DR EMBL: U40278; AAB60638.1; JOINED.  
 DR EMBL: AF187881; AAF62875.1; -;  
 DR HSSP: P11215; 1BHQ.  
 DR Genew; HGNC:6146; ITGAD.  
 DR MIM; 602453; -;  
 DR GO; GO:0008305; C:integrin complex; TAS.  
 DR GO; GO:0016337; P:cell-cell adhesion; NAS.  
 DR GO; GO:0007160; P:cell-matrix adhesion; NAS.  
 DR GO; GO:0006955; P:immune response; NAS.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; WFA; 1.  
 DR PRINTS; PR01185; INTEGRIN.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; WFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWFA; 1.  
 KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 17 Potential.

CHAIN	18	1162	Integrin alpha-D.
DOMAIN	18	1100	Extracellular (Potential).
TRANSMEM	1101	1121	Potential.
DOMAIN	1122	1162	Cytoplasmic (Potential).
REPEAT	32	85	FG-GAP 1.
REPEAT	86	?	FG-GAP 2.
DOMAIN	150	332	VWFA.
REPEAT	350	400	FG-GAP 3.
REPEAT	401	452	FG-GAP 4.
REPEAT	454	516	FG-GAP 5.
REPEAT	518	576	FG-GAP 6.
REPEAT	581	633	FG-GAP 7.
CA_BIND	465	473	Potential.
CA_BIND	530	538	Potential.
CA_BIND	593	601	Potential.
SITE	1127	1131	GFPR motif.
DISULFID	67	74	By similarity.
DISULFID	106	124	By similarity.
DISULFID	655	710	By similarity.
DISULFID	769	775	By similarity.
DISULFID	846	861	By similarity.
DISULFID	994	1018	By similarity.
DISULFID	1023	1028	By similarity.
CARBOHYD	59	59	N-linked (GlcNAc. .) (Potential).
CARBOHYD	87	87	N-linked (GlcNAc. .) (Potential).
CARBOHYD	99	99	N-linked (GlcNAc. .) (Potential).
CARBOHYD	391	391	N-linked (GlcNAc. .) (Potential).
CARBOHYD	691	691	N-linked (GlcNAc. .) (Potential).
CARBOHYD	733	733	N-linked (GlcNAc. .) (Potential).
CARBOHYD	873	873	N-linked (GlcNAc. .) (Potential).
CARBOHYD	957	957	N-linked (GlcNAc. .) (Potential).
CARBOHYD	1046	1046	N-linked (GlcNAc. .) (Potential).
CONFLICT	500	500	Missing (in Ref. 2).
CONFLICT	515	518	GHFW -> ATP (in Ref. 2).
CONFLICT	825	825	L -> V (in Ref. 2).
CONFLICT	984	984	V -> A (in Ref. 2).
SEQUENCE	1162 AA;	126885 MW;	F296A1A35455D77D CRC64;

Query Match 58.2%; Score 3417; DB 1; Length 1162;  
 Best Local Similarity 59.8%; Pred. No. 3.7e-217;  
 Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

QY	1	ENLDTNMTFQSNARGFGQSVVQLQGSRRVVGAPQEIIVAAVNRGSLYOCYDSTGSCPI	60
Db	18	FNLDVEEPTIFQEDAGGFGQSVVQFGSRVVGAPLEVAANOTGRLYDCAAATGMCOPI	77
QY	61	RLQVPVEAVNMSLGLSLAATTSPQLACGPTVHQTCSNTYVYKGLCFLFGSNLRQOPK	120
Db	78	PLHIRPEAVNMSLGLTLAASTNGSRLLACGPTLHRVCGENSYSKSGCLLGSRW-ETIQT	136
QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTKTFLSLMOYSEEF	180
Db	137	VDPATPECPHQEMDVIIVFLIDGSGSIDONDFNQMKGFAVQVWGQFEGDTLFLALMOYSNLL	196
QY	181	RIHFTPEFQNNPNRSLVKPTQLLGRTHRTATGVKVKRELINITNGARKNAFKILIVI	240
Db	197	KIHFTFTQFTSPSQSLVDPIVLQKGLTFTATGILTVTLFHHKNGARKSAKKILIVI	256
QY	241	TGKEFGDPLGYEDVTPEDADREGVIRYIVGVGDAPFSEKSRQBELNTIASKPRDHFQVN	300
Db	257	TDQQKYKDPLEYSVDVIPQAEKAGIIRYAIQVGHAFQGPATQBELNTISSAPPDHFQKVD	316
QY	301	NFEALKTIQNLREKIFAECTQTSSSSFEHMSQEGSAITSNGPLSTVGSVDWAG	360
Db	317	NFAALGSIQQLQEKIYAVEGTQSRASSQFHEMSQEGFSTALTMDFLGAVGFSWSG	376
QY	361	GVFLYTSKEKSTFINNTRVDSDDNDAYLGAAYAILNRNVSQSLVGLAPRYOHLGLVAMFR	420
Db	377	GAFLYPNMSPFTINMSQENVMDRDSYLGSTELALWGVQNLVGLAPRYOHTGKAVIFT	436
QY	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEOTRGQSVCPPL	480
Db	437	QVSRQWRKKAETVGTQIGSVFGASLCSVDVDSGSTDLLILIGAPHYVEOTRGQSVCPPL	496

```
QY 481 PRGORARWQCDVAVLYGEGQGPWGFRGAALTIVLGDVNGDKLTDVAIGARBEDNRGAVYLF 540
Db 497 ERGORVQWQCDVAVLYGEGQGPWGFRGAALTIVLGDVNGDKLTDVAIGARBEDNRGAVYLF 556
QY 541 HGTSGSGISPSHRSORIASGKLSPLRQYFGQSLGGQDLTMDGLVDLTWGAQGHVLLLRSQ 600
Db 557 HGASESGISPSHRSORIASGKLSPLRQYFGQSLGGQDLTMDGLVDLWGAQGHVLLLRSL 616
QY 601 PVLRVKAIMENPREVARNPEFCNDQVVKYKGEAEVRVCLHVQKSTRDRILREGQIQSVVT 660
Db 617 PVLKVGVMARFSPVEKAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSEVR 674
QY 661 YDLALDGRPHSRHRAVENETKNSRROTQVGLGTCTETKLQLOPNCIEDPSPVILRLNF 720
Db 675 FDLALDGRPHSRHRAVENETKNSRROTQVGLGTCTETKLQLOPNCIEDPSPVILRLNF 734
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 735 SLVREPIPSQNLRLPVLAVGSDQLFTASLPPEKNCQDGLCEGLGVTLSPFGLQTLTVG 794
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYKSVSTLQNRQSRQWRLACBSASTEV 840
Db 795 SSELNVITVVRNDGEDSYRTQVTFPPPLDLSYKSVSTLQNRQSRQWRLACBSASTEV 853
QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900
Db 854 EG-LRSSRCSVNHPIFEGSNGTFTVTVDSYKATLGDRLMRASASNNKASSKATP 912
QY 901 QLELPVKYAVVMVYTSHGVSSTKYNLF-TASENTSRVMQHOYQVSNLQORSIPISLVLFP 959
Db 913 QLELPVKYAVVMVYTSHGVSSTKYNLF-TASENTSRVMQHOYQVSNLQORSIPISLVLFP 972
QY 960 VRLNQTVMDRPPQVTFSENLSSTCTYERLPDSDFLAELRKAPVNCSTIACVORIQCDI 1019
Db 973 VLLNGVAVMDVMEAPSQSL--PCVSRKPKPQHSDFLTQTSRSPMLPCSIADCLQFRCDV 1030
QY 1020 PFGIQEENFATLKNLSFDWYIKTSHNHLIVSTAELFENDSVFTLLPGQAFVRSOTE 1079
Db 1031 PFSVSQELDFTLKGNLSPGWRTLOKKVLIVSVAEITFDTSVYSQLPQGEAFMRAQME 1090
QY 1080 TKVEPFVPPNPLPLIVGSSVGGLLALLIITAALYKLGFFKRYKXDMSE 1128
Db 1091 MWLEEDVNAIPIIMSSVGCALLLALITATLYKLGFFKRYKXDMSE 1139

RESULT 7
Q6KAS4 PRELIMINARY; PRT; 1188 AA.
AC Q6KAS4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE MFLJ00114 protein (fragment).
GN Name=MFLJ00114;
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kituno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.,
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Genes:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-Homologous cDNAs
RT Identified by Screening of Terminal Sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries."
RL DNA Res. 11:167-180(2004).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
DR EMBL; AK311133; BAD21383.1; -
DR GO; GO:0009897; C:external side of plasma membrane; IDA.
DR InterPro; IPR000413; Integrin_alpha.
```

```
InterPro; IPR002035; VWF_A.
Pfam; PF01839; FG-GAP; 3.
Pfam; PF00357; Integrin_alpha; 1.
Pfam; PF00092; VWA; 1.
PRINTS; PR01185; INTEGRIN.
PRINTS; PR00453; VWFADOMAIN.
SMART; SM00191; Int_alpha; 5.
SMART; SM00327; VWA; 1.
PROSITE; PS00242; INTEGRIN_ALPHA; UNKNOWN_1.
PROSITE; PS02334; VWFA; 1.
KW Cell adhesion; Integrin; Transmembrane.
FT NON TER 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.8%; Score 3335.5; DB 2; Length 1188;
Best Local Similarity 57.4%; Pred. No. 9.8e-212;
Matches 654; Conservative 165; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFQENARFGQSVVQLOQSGRVVVGAPQIIVAAQRGSLYQCDYSTGSCBPI 60
Db 39 FNLDAREPPTHFMDCAEFCHSVIQQDSSVVVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 98
QY 61 RLQVPEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYKGLCFELGSLNRQOPQK 120
Db 99 SLQVPEAVNMSLGLSLAATNPSWLLACGPTVHTTCRENIYLTGLCFLLSSSFQKS-QN 157
QY 121 FPEALRCPOEDSDIAFLIDGSGSIIPHDPRMKKEFVSTWMEQLKKSKTILFSLMQSEEF 180
Db 158 FPTAAQCPQDQDIVFLIDGSGSISSTDFEKLDFKAVMSQLQRESTRFSLMQFSDYF 217
QY 181 RIHTFTEFQNNPRSLVPIPTOLLGRTHATGVKRVIRELLNITMGARKNAFKILIVI 240
Db 218 RVHTFNFTSPLSLDLSVRQLRGYTYTASAKHIVITELFTTQSGARQDATKVLIVI 277
QY 241 TDGEKFGDPIGYEDVIEPADREGVIRYVIGVDAPRSEKSRQELNITASKPRDHVQVN 300
Db 278 TDRKQKQDNLSDYSVIPMAEAAAIIRVAIGVKAFYNEHSKQELKALIASMPSEHYVFSVE 337
QY 301 NFPAKTIQNLREKIPAIETGTGSSSSPEHMSQEGFSAATITSGNPLISTVCSYDWAG 360
Db 338 NFDAKDIEQLKEKIPAIETGTGSSSTFELEMSQEGFSAVTPDGPVLGAVGFSWSG 397
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDIYVAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420
Db 398 GAFLYPSNMPTFINMSQENEDMRDVLGYSTALAFWKGVHSLILGAPRHQHTKGVIFT 457
QY 421 QNTGMMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYTYEQTRGGQVSCPL 480
Db 458 QESRHRPKSEVRGTQIGSYFGASLCSVDMDRDGSTDVLIGVPHYTYEHTRGQVSCPM 517
QY 481 PRGORARWQCDVAVLYGEGQGPWGFRGAALTIVLGDVNGDKLTDVAIGARBEDNRGAVYLF 540
Db 518 P-GVGRRWHCQGTTLHGQGHQWFRFGAALTIVLGDVNGDSLADVAIGAAGPEENRGAVYIF 576
QY 541 HGTSGSGISPSHRSORIASGKLSPLRQYFGQSLGGQDLTMDGLVDLTWGAQGHVLLLRSQ 600
Db 577 HGASRQDIAPSPQRIASQIPRIQYFGQSLGGQDLTMDGLVDLWGAQGHVLLLRTR 636
QY 601 PVLRVKAIMENPREVARNPEFCNDQVVKYKGEAEVRVCLHVQKSTRDRILREGQIQSVVT 660
Db 637 PILRVSPVTHFTPAETISRSVFCEQVAPETQLSDATVCLHITHESPKTQL--GDLRSTVT 694
QY 661 YDLALDGRPHSRHRAVENETKNSRROTQVGLGTCTETKLQLOPNCIEDPSPVILRLNF 720
Db 695 FDLALDGRHGRUSTRAIKFKETKTRALTRVTKLGNKHCESVKLLLPACVEDSVTPTILRLNF 754
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFALFPPEKNCNDNICODDLSITFSFMSLDCLVVG 780
Db 755 SLVGVPISSIQNLQPMPLAVDQTYFTASLPPEKNCADHICQDDLSVVFPGPDLKTLVVG 814
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYKSVSTLQ-----NQSRQSWR 829
Db 815 SDLELVNDVTVSNDGEDSYGTITVTFYFVGLSFRVAEGQVFLRKKEKDQWQRKGQSLH 874
```

```

QY 830 LACESASSTEVSGALKSTSCSINHPIFPENSEVTNITFDVDSKASIGNKLLKANVTSE 889
DB 875 LMCD--STDRSQGLWSTSCSRHVI FRGGSQMTFLVTFDVSFKABLGDRLLRLARVGSE 932
QY 890 NNMPRTNKTEFOLELPVKVAVYVTVSHGVSTKYNLFTASE-NTSRVMOHOVQVSNLQOR 948
DB 933 NNVPCTPKTTFQLELPVKVAVYVTVSHGVSTKYNLFTASE-NTSRVMOHOVQVSNLQOR 992
QY 949 SUPISLVFLVPVRLNQTVDWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCS 1008
DB 993 DVPVSINFVWPIELKGEAVW-TVMVSHPNQPLTQCYRNLKPTQDFLTHMQSPVLDGS 1051
QY 1009 IAVCQRIQCDIPFFGIQEEFNATLGNLSFDMYIKTSHNHLIVSTAELFNDVSFTLLP 1068
DB 1052 IADCLHLRCDIPSLGILDLYFLKGNLSFGWISQTLQKKVLLSEAEITFNTSVYSQLP 1111
QY 1069 GQAFVRSOTETKVEPFEPVNPPLIVGSSVGGLLALITAAIYKLGFFKROKYMMSSE 1128
DB 1112 GQAEFLRAQTKTVLEMYKVHNPVPLIVGSSVGGLLALITAAIYKAGFFKROKYMLEE 1171

RESULT 8
BAD21383
ID BAD21383 PRELIMINARY; PRT; 1188 AA.
AC BAD21383;
DT 01-JUN-2004 (T-EMBLrel. 27, Created)
DT 01-JUN-2004 (T-EMBLrel. 27, Last sequence update)
DE MFLJ00114 protein (Fragment).
GN MFLJ00114.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Okazaki N., Kikuno R., Ohara R., Inamoto S., Koseki H., Hiraoka S.,
RA Saga Y., Kitamura H., Nakagawa T., Nagase T., Ohara O., Koga H.;
RT "Prediction of the Coding Sequences of Mouse Homologues of FLJ Gene:
RT The Complete Nucleotide Sequences of 110 Mouse FLJ-homologous cDNAs
RT Identified by Screening of Terminal sequences of cDNA Clones Randomly
RT Sampled from Size-Fractionated Libraries.";
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL, AKL31133; BAD21383.1; -.
FT NON_TER 1
SQ SEQUENCE 1188 AA; 131248 MW; B8D93107BDBB4178 CRC64;

Query Match 56.8%; Score 3335.5; DB 2; Length 1188;
Best Local Similarity 57.4%; Pred. No. 9.8e-212;
Matches 654; Conservative 165; Mismatches 302; Indels 19; Gaps 7;

QY 1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVGAPQEIIVAAVQVSLYQCYDSTGSCPEI 60
DB 39 FNLDAEKPTFHFDGABFGHVSQYDSSWVVGAPKEIKATNQIGGLYKCGYHTGNCEPI 98
QY 61 RLQVPVEAVNMSLGLSLAATSTPQLLACGPTVHTQCTSENYVKGICFLFGSNLRQOPQK 120
DB 99 SLQVPPEAVNMSLGLSLAATSTPQLLACGPTVHTQCTSENYVKGICFLFGSNLRQOPQK 157
QY 121 FPEALGCPQEDSDIAFLIDGSGIIPHDPRRMEKFFVSTVMEQLKKSKTLFSLMOMYSEEF 180
DB 158 FPTAQECPCQODIVFLIDGSGISSTDFEKMLDFVKVAMSQLQRPSTFSLMQSDYF 217
QY 181 RIHFTFKFQNNPRSLVKPITQLLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 240
DB 218 RVHFTFNFTSTSSPLSDSVQLRGYTYTASAIKHVITELFTTQSGARQDATKVLIVI 277
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVGAFAFSEKRSQELNITIASPPRDHVFQVN 300
DB 278 TDGRKQGDNLSDYSDVIPMAEAAIIRYVIGVGAFAFSEKRSQELNITIASPPRDHVFQVN 337
QY 301 NFEALTKIQNLREKIFAISGTQTGSSSSFEHMSQEGFSAAITSNGLPLSTVGSYDWAG 360

```

## RESULT 9

```

ITAX_MOUSE
ID ITAX_MOUSE STANDARD; PRT; 1169 AA.
AC Q9QXH4;
DT 05-JUL-2004 (Rel. 44, Created)
DT 05-JUL-2004 (Rel. 44, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Integrin alpha-X precursor (leukocyte adhesion glycoprotein p150,95
DE alpha chain) (leukocyte adhesion receptor p150,95) (CD11c).
GN Name=Itgax;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dendritic cell;

```

RA Huang X., Gorski K., Tong C., Rattis F.-M., Teeng S.-Y., Pardoll D.,  
 RA Tsuchiya H.;  
 RL "Isolation of genes selectively expressed by dendritic cells.";  
 CC Submitted (DSC-1999) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Integrin alpha-X/beta-2 is a receptor for fibrinogen. It  
 CC recognizes the sequence G-P-R in fibrinogen. It mediates cell-cell  
 CC interaction during inflammatory responses. It is especially  
 CC important in monocyte adhesion and chemotaxis (By similarity).  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-X  
 CC associates with beta-2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF211864; AAP23492.1; -;  
 DR HSSP; P20702; IN3Y.  
 DR MGI; MGI:96609; Itgax.  
 DR GO; GO:0009897; C:external side of plasma membrane; IDA.  
 DR InterPro; IPR00413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWF; 1.  
 DR PRINTS; PRO1185; INTEGRIN.  
 DR PRINTS; PRO0453; VWFADOMAIN.  
 DR SMART; SMO0191; Int alpha; 5.  
 DR SMART; SMO0327; VWF; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS50234; VWFA; 1.  
 DR Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 19 Potential.  
 FT CHAIN 20 1169 Integrin alpha-X.  
 FT DOMAIN 20 1116 Extracellular (Potential).  
 FT TRANSMEM 1117 1137 Potential.  
 FT DOMAIN 1138 1169 Cytoplasmic (Potential).  
 FT REPEAT 34 87 FG-GAP 1.  
 FT REPEAT 88 ? FG-GAP 2.  
 FT DOMAIN 152 330 VWFA.  
 FT REPEAT ? 402 FG-GAP 3.  
 FT REPEAT 403 454 FG-GAP 4.  
 FT REPEAT 456 518 FG-GAP 5.  
 FT REPEAT 519 577 FG-GAP 6.  
 FT REPEAT 582 634 FG-GAP 7.  
 FT CA BIND 467 475 Potential.  
 FT CA BIND 531 539 Potential.  
 FT CA BIND 594 602 Potential.  
 FT SITE 1140 1144 GFFKR motif.  
 FT DISULFID 69 76 By similarity.  
 FT DISULFID 108 126 By similarity.  
 FT DISULFID 656 711 By similarity.  
 FT DISULFID 770 776 By similarity.  
 FT DISULFID 858 873 By similarity.  
 FT DISULFID 1007 1031 By similarity.  
 FT DISULFID 1036 1041 By similarity.  
 FT CARBOHYD 89 89 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 267 267 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 393 393 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 734 734 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 949 949 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 1059 1059 N-linked (GlcNAc... ) (Potential).  
 FT CARBOHYD 1084 1084 N-linked (GlcNAc... ) (Potential).

SQ SEQUENCE 1169 AA; 129150 MW; C6164120332C19A6 CRC64;  
 Query Match 56.6%; Score 3326.5; DB 1; Length 1169;  
 Best Local Similarity 57.2%; Pred. No. 3.8e-211;  
 Matches 652; Conservative 167; Mismatches 302; Indels 19; Gaps 7;  
 Qy 1 FNLDTNAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVANORGSLYQCDYSTGSCPEI 60  
 Db 20 FNLDAEKLTHFMDGAEFGHSLVLYDSSVVVVGAPKEIKATNQIGGKYKGYHTGNCPEI 79  
 Qy 61 RLQVPVEAVNMSLGLSLAAATSPPLQACAGPTVHQTCSNTYVKGCLFLFGSNLRQOPQK 120  
 Db 80 SLOVPPPAVNISLGLSLAAATNPSPWLLACGPTVHTCRENIYLTGLCFLLSSSKQS-QN 138  
 Qy 121 FPEALRCGPQEDSDIAFLIDGSGSIIPHDFRMRKEFVSTVMEQKKSKTLPESLMQYSEEF 180  
 Db 139 FPTAQACPKQDQDIVELIDGSGSISSTDFEKMDFVKAVMSQLQRPSTRESLMQFSDYF 198  
 Qy 181 RIHFTFKFONNPNRSLVAPITOLLGKTHATGVKRVIRELLMTNGARKNAFKILIVI 240  
 Db 199 RVHFTFNFTSSEPLSLGSRQLRGTYTASAKHIVITELFTTQSGARQDARKVLIVI 258  
 Qy 241 TDEKFGDPLGYEDVIEADREGVIRYVIGVDARFSEKSRQELNTIASKEPPRHVFOVN 300  
 Db 259 TDCRKQGDNLSDYSVIPMAEAAAIIRVAIGVKAFYNEHSKQELKAIASMPSEHYVFSVE 318  
 Qy 301 NFEALKTIQNLREKIPAIETGTGTGSSSSSEHEMSQGFSAITNSNPLSTVGSYDWAG 360  
 Db 319 NFDALKDIENQLKEKIPAIETGTETPSSSTFELEMSQGFSAVFTPDGVLGAVGSFWSG 378  
 Qy 361 GVFLYTSKEKSTFINMTFRVSDMNDAYLGYAAAILIILNRVQSLVILGAPRYOHIGLVAMER 420  
 Db 379 GAFLYPSNMPTPTINMSQENEDMRDALVLTALAFWKGVHSLILGAPRHQHTOKVIFT 438  
 Qy 421 QNTGMWESNANVKGITQIGAYFGASLCSVDVDSNGSTDLVILGAPHYEYQTRGGQVSCPL 480  
 Db 439 QESRHWPKGEVRGTQIGSYFGASLCSVDMDRDGSDTLVLIGVPHYEYTRGGQVSCPM 498  
 Qy 481 PRGQARWQCDVLYGEGQOPWGEFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVLYF 540  
 Db 499 P-GVGRHWCHGTTLHGEGHPWGRFGAALTVLGDVNGDSLADVAIGAAGPEENRGAVYIF 557  
 Qy 541 HGTSGSGISPSHSORAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAGQHVLLLRSQ 600  
 Db 558 HGASRQDIAPSPQISASQIPRIQYFGQSLGGQDLTRDGLVDLAVGSKGRVLLLR 617  
 Qy 601 FVLRVKAIMEFPNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIQSVVT 660  
 Db 618 PILRVSPVHTPTAEISRSVFCEQEQVAPETLLSDATVCLHIHESPKTQL--GDLRSTVT 675  
 Qy 661 YDLALDSGRPHSRAVENETKSTRTQTVLGLTQTCETLKLQLPNCIEDPVSPIVLINP 720  
 Db 676 FDALADHGRSLSTRAIFKETKTRALT RVKTLGLKNKCESVKLLLPACVEDSVPTITLRF 735  
 Qy 721 SLVGTPLSAFQNLRLPVLAEADAQRLFTALFPFKKNCNDNI CQDDLSTFTFSMDCLVVG 780  
 Db 736 SLVGVPISSLQNLQPLAVDDQYFTASLPFKKNCADHICQDDL SVVFGFDPDKTLVVG 795  
 Qy 781 GPREFNVTVTRNDGEDSYETQVTFPPPLDLSVRKYSTLQ-----NQRSQSRWR 829  
 Db 796 SDLELNVDVTVSNDGEDSYGTVTTLFYPVGLSFRRAEGQVFLRKKEKEDQWQRGQSHLH 855  
 Qy 830 LACESASSTEVSGALKSTCSINHPFPENSEVENTITFDVDSKASIGNKLLKANKYTS 889  
 Db 856 LMCD--STPRSOGLWSTSCSRHVIFRGSGQMTFLVTFDVSRAELGDLLRLRARVGE 913  
 Qy 890 NNPRTNKTEFQLELPVKYAVVMVYSHGVSTKYLNFTASE--NTSRVMQHQYQVSNLQGR 948  
 Db 914 NNVPGPPTKTFQLELPVKYAVYTWISSHDQFTKYLNFTSEKETSVEVHRFQVNNLQGR 973  
 Qy 949 SLPISLVFLVPVRLNQTVIWRPQVTFSENLSSCTCHTKERLPSHSDFLAELRKAPVNCVS 1008  
 Db 974 DVPVSINFVPIELKGEAVW--TVMVSHPQNPLTQCYRNRLKPTQFDLLTHMOKSPVLDGS 1032



QY 1009 IAVCORIQCDIPFGIOEBENATLKGNSFDWIKTSHNELLIVSTAEILLNDVSFTLLP 1068  
 Db 1033 IADCLHRCIDPSGLDELDELFLKGNLSFGWISQTLQKXVLLSBAEITNTSVYSQLP 1092  
 QY 1069 GQAGFVRSQTETKVEPEFVENPLPLVSGVGLLALITAAIYKLGFFKROKDMSE 1128  
 Db 1093 GQAFATRAQTKVLEMYKHNPVPLVSGVGLLALITAILYKAGFFKROKEMLEE 1152

RESULT 10  
 ITAD RAT  
 ID ITAD RAT STANDARD; PRT; 1161 AA.  
 AC OSQVE7;  
 DT 05-JUL-2004 (Rel. 44, Created)  
 DT 05-JUL-2004 (Rel. 44, Last sequence update)  
 DE Integrin alpha-D precursor.  
 GN Name=Itgad;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RC STRAIN=Sprague-Dawley;  
 RA O'Brien M.M., VanderVieren M., Kilgannon P.D., Dietsch G.,  
 RA Gallatin W.M.;  
 RT "Cloning of rat alpha D, a novel beta 2 integrin.";  
 RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Integrin alpha-D/beta-2 is a receptor for ICAM3 and  
 CC VCAM1. May play a role in the atherosclerotic process such as  
 CC clearing lipoproteins from plaques and in phagocytosis of blood-  
 CC borne pathogens, particulate matter, and senescent erythrocytes  
 CC from the blood (By similarity).  
 CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-D  
 CC associates with beta-2 (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins  
 CC with I-domains do not undergo protease cleavage.  
 CC -!- SIMILARITY: Belongs to the integrin alpha chain family.  
 CC -!- SIMILARITY: Contains 7 FG-GAP repeats.  
 CC -!- SIMILARITY: Contains 1 VWFA domain.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AF021334; AAF21241.1; -;  
 DR HSSP; P11215; 1BHQ.  
 DR InterPro; IPR000413; Integrin\_alpha.  
 DR InterPro; IPR002035; VWFA\_A.  
 DR Pfam; PF01839; FG-GAP; 3.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWFA; 1.  
 DR PRINTS; PR01185; INTEGRINA.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWFA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS50234; VWFA; 1.  
 KW Calcium; Cell adhesion; Glycoprotein; Integrin; Magnesium; Receptor;  
 KW Repeat; Signal; Transmembrane.  
 FT SIGNAL 1 19 Potential.  
 FT CHAIN 20 1161 Integrin alpha-D.  
 FT DOMAIN 20 1100 Extracellular (Potential).  
 FT TRANSMEM 1101 1121 Potential.  
 FT DOMAIN 1122 1161 Cytoplasmic (Potential).  
 FT REPEAT 34 87 FG-GAP 1.

FT REPEAT 88 ? FG-GAP 2.  
 FT DOMAIN 152 334 VWFA.  
 FT REPEAT 352 402 FG-GAP 3.  
 FT REPEAT 403 454 FG-GAP 4.  
 FT REPEAT 456 517 FG-GAP 5.  
 FT REPEAT 519 577 FG-GAP 6.  
 FT REPEAT 582 634 FG-GAP 7.  
 FT CA\_BIND 467 475 Potential.  
 FT CA\_BIND 531 539 Potential.  
 FT CA\_BIND 594 602 Potential.  
 FT SITE 1126 1130 GFKR motif.  
 FT DISULFID 69 76 By similarity.  
 FT DISULFID 108 126 By similarity.  
 FT DISULFID 656 711 By similarity.  
 FT DISULFID 769 775 By similarity.  
 FT DISULFID 845 860 By similarity.  
 FT DISULFID 993 1017 By similarity.  
 FT DISULFID 1022 1027 By similarity.  
 FT CARBOHYD 61 61 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 245 245 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 393 393 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 696 696 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 734 734 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 784 784 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 907 907 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 936 936 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 1045 1045 N-linked (GlcNAc...) (Potential).  
 SQ SEQUENCE 1161 AA; 126600 MW; 2258491A984705E CRC64;

Query Match 55.3%; Score 3251.5; DB 1; Length 1161;  
 Best Local Similarity 58.0%; Pred. No. 3.5e-206;  
 Matches 634; Conservative 158; Mismatches 302; Indels 13; Gaps 9;

QY 2 NLDNTAMTFQENARGFGQSVVOLQSGRVVVGAPQEIIVAAANQRGSLYQCDYSTGSCPIR 61  
 Db 21 NLDVVEPIVREDAAASFGTVVQFGGSLVVGAPLEAVAVNOTGRDLYDCAPATGMCPIV 80  
 QY 62 LQVPVEAVNMSGLSLAATSPDQLACGPTVHQTCTSENTYVKGCLFQSNLRQDQKF 121  
 Db 81 LRSPLAVNMSGLSLVATATNNAQLACGPTAQRACVKNMYAKSGCLLLGSSL-QFTQAV 139  
 QY 122 PEALRGCPQEDSDIAFLIDGSGSIIIPDPRMKFEVSTWEOQLKSKTLPSLMQYSEER 181  
 Db 140 PASMPCPRQEMDIALIDGSGSINORDFAQKADFKVLMGEFASSTLPSLMQYSNLK 199  
 QY 182 IHPTTFEFQNNPNRSLVKPITQLGRTHATGVKRVIRELLNITNGARKNAFKILIVIT 241  
 Db 200 THPTTFEFKNILDPQSLVDPIVQLQGLTYTATGIRTVMEELFHSKNGSRKSAKKILLVIT 259  
 QY 242 DGEKFGDPLGYEDVPEADREGVIRVIGVDAPFRSEKSRQELNTTASKPPRDHVFQVNN 301  
 Db 260 DGQKRDPLEYSDVIPAADKAGIIRYAIQVGFQDAFOEPTALKELNTTIGSAPPDQHVFKVGN 319  
 QY 302 FEALKTIQNLREKIFAIEGTOTGSSSSFEHMSQEGFSAATITNSGPLLSTVGSYDWAGG 361  
 Db 320 FAALRSIQQLQKEKIFAIEGTQSRSSSFQHEMSQEGFSALTSDGPVLGAVGSFWSGG 379  
 QY 362 VFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILRNRRVQSLVGLGAPRYOHIGLVAMFRQ 421  
 Db 380 AFLYPNTRPTFINMSQENVMDRDSYLGYSTAVAFWKGVHSLILGAPRHQHTGKVIFTQ 439  
 QY 422 NTGMWESNANVKTQIGAYFGASCLSDVDVDSNGSTDLVLIGAPHYEQTRGGQVSCPLP 481  
 Db 440 EARHWRPKSEVRGTQIGSYFGASCLSDVDVDRDGSSTDLVLIGAPHYEQTRGGQVSPVPV 499  
 QY 482 RQORAWQCDVLYGQGGPQWGFAGALTGLGVNGDKLTDVAIGAPGEEDNEGAVYLFH 541  
 Db 500 -GVRGEWQCEATLHGQGHFWGFVGVNADVAIGAPGEESRGAVYIFH 558  
 QY 542 GTSGSGISPSHSORIASKLSPRLQYFGQSLSGQDLTMDGLVDLTVAGQGHVLLRSQP 601  
 Db 559 GASRLIMPSPSRVITGSQLSLQYFGQSLSGQDLTQDGLVDLTVAGQGHVLLRSUP 618



```
QY 239 VITDGRKFDPLGYEDVIEADREGRVIRVGVGDAPRSEKSRQELNTIASKPRPHVFQ 238
Db 258 IITDGEATD-----ENIDDA--KQIRIYIGKFKTKESQEAHQFASKVEBPFVKI 310
QY 299 VNNFEALKTIQNLREKIFAIEGTQTGSSSFHEMSQEGFSAITSGNPLLTGVSYDW 358
Db 311 LDTFEKLDKDLFTLEQKIVIEGTQKDLTSFNWELSSGSIADLSGEGHVAVGAKDW 370
QY 359 AGGVF-LYTSKEKSTFNMTVRDSDMDNDAYLGAFAA-IILNRVQSILVGLGAPYQHIGLV 416
Db 371 AGGFLDKADLKSTTVGNEPLTVESRAGYLVTVTWLPKRGMSLLIATGAPYQHVGRV 430
QY 417 AMFRO--NTGMWESNAVKTQIGVAFGASLCSVDVDSNGSTDLVLIGAPHYEQTGGQ 474
Db 431 LUFQPKRGPMWSQIQEIDIGIGSTFGGELCGVDVDRGETELLIAAPLYGEORGR 490
QY 475 VSYCPFRGORARWQCDVLYGEGQPGWGRFGAALTVLGVDNGDKLTDVAIGAPGEBEDNR 534
Db 491 VFYI---QKIQLEFQMVSELQGETGYPLGRFGAIAALTIDNGDELTDVAVGAPLEE--Q 545
QY 535 GAVYLFHGTSGSISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTWDLGLVDLTGAGQHV 594
Db 546 GAVYIFNGQGG--GLSPRPSQRIEQTQMFSGIQWFGRSIHGVKDLGGDLADVAVGAEGQV 604
QY 595 LLRSOPVLVRKAIMFENPREVARNVFECDNVVKGKEAG-EVRVCLHVQK--STRDLR 651
Db 605 IVLSSRPVDIITVSFSPAPIEPVHEVECSYSTSNQKKEGNNLTVCQVKSLLST----F 660
QY 652 EGQIQSVVTVYDLALDGRPHSRVAFNETKNSTRQTQVLGLTQTCETKLQLPNCIEDPV 711
Db 661 QGHLVANLTVTLQDGHRTSGLFGPGKHKLIGNTAVTPV-KSCFVFWFHFICIQDLI 719
QY 712 SPIVLIRNFSL---VGTPLS--AFGNLRPVLAEDAQLFTALPPFKNCGNDNICODDLS 766
Db 720 SPINVSLSYSLWEEETPRDPRALDRDIPPLKPSPHLETKEIPFEKNCGEDKNCEADLK 779
QY 767 ITFSFMSLCLVVGPPREFNVTVVRNDGDSYRTQVTFPEPLDLSYRKVSTILQONRSOR 826
Db 780 LAFSDMRSKILRUTPSSASLSVRUTLRNTAEDATWVQVTLSPGGLSFRKVEIL---KPHS 836
QY 827 SWRLACESASTVBVGALKSTCSINHPIFENSEVTFNITFDVDSKASLGNKLLKANV 886
Db 837 HVPVGCHELPEEAVVHS-RALSCNVSSPIFGESMWDIQWENTLQKSGWGDFTLQANV 895
QY 887 TS-----ENNMPRTKTEFQLELPVKYAVVMVTVSHGVSTKYLNTASENTSRVMQH 938
Db 896 SCNNEDSLLLEDNSATTS-----IPVMYPINVLTKDQENSTLYISFTPKSPKIHVKH 948
QY 939 QYQV----SNLGRSLPISLVLPVRLNOTVI---WD---RQVTFPS-ENLSSTCHKE 987
Db 949 IYQVRQPSNYDMP-PLEALVRVRVHSEGLTHKWSIQMEPPVNCSPRNLESPSDEAE 1007
QY 988 RLPSSHDFLAELRKAPVYVNCIAVCQRIQCDIPFGIQEEFNATLKGNSLFDWYIKTSHN 1047
Db 1008 -----SCSFGT--EFRCPIDF---RQELIVQVNGMVELRGFTIKAS-S 1043
QY 1048 HLLIVSTAEILFNDSVFTLPQGAQFVRSQTEKVFEPFVNPPLPLIVGSSVGLLILLAL 1107
Db 1044 MSLCSLSLASFNSKHFHLYGRNASM-AQVMVKVDLVYKEMLLYLVLSIGIGLGLLLFL 1102
QY 1108 ITAALYKLGFFKQYKDMV-----SEGGPPGAPQ 1137
Db 1103 IFALYKVGFFKRLKEMNEAVDASSEIFGEDAGQPELEKE 1144
```

## RESULT 12

```
ITAL_HUMAN STANDARD; PRT; 1170 AA.
AC P20701; O43746;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1
```

```
DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha
DE chain) (CD11a).
GN Names:ITGAL; Synonyms:CD11a;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
RX MEDLINE=89139587; PubMed=2537322;
RA Larson R.S., Corbi A.L., Berman L., Springer T.;
RT "Primary structure of the leukocyte function-associated molecule-1
RT alpha subunit: an integrin with an embedded domain defining a protein
RT superfamily.";
RL J. Cell Biol. 108:703-712(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99425270; PubMed=10493829;
RA Loftus B.J., Kim U.-J., Sneddon V.P., Kalush F., Brandon R.,
RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
RA Deslattes Ways A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
RT "Genome duplications and other features in 12 Mb of DNA sequence from
RT human chromosome 16p and 16q.";
RL Genomics 60:295-308(1999).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 153-335, AND REVISION TO 214.
RX MEDLINE=96036087; PubMed=7479767;
RA Qu A., Leahy D.J.;
RT "Crystal structure of the I-domain from the CD11a/CD18 (LFA-1, alpha L
RT beta 2) integrin.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:10277-10281(1995).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 153-335.
RX MEDLINE=96398682; PubMed=8805579;
RA Qu A., Leahy D.J.;
RT "The role of the divalent cation in the structure of the I domain from
RT the CD11a/CD18 integrin.";
RL Structure 4:931-942(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 153-334.
RX MEDLINE=99425288; PubMed=10493852;
RA Kallen J., Welzenbach K., Ramage P., Geyl D., Kriwacki R., Legge G.,
RA Cottens S., Weitz-Schmidt G., Hommel U.;
RT "Structural basis for LFA-1 inhibition upon lovaastatin binding to the
RT CD11a I-domain.";
RL J. Mol. Biol. 292:1-9(1999).
CC -!- FUNCTION: Integrin alpha-L/beta-2 is a receptor for ICAM1, ICAM2,
CC ICAM3 and ICAM4. It is involved in a variety of immune phenomena
CC including leukocyte-endothelial cell interaction, cytotoxic T-cell
CC mediated killing, and antibody dependent killing by granulocytes
CC and monocytes.
CC -!- SUBUNIT: Heterodimer of an alpha and a beta subunit. Alpha-L
CC associates with beta-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P20701-1; Sequence=Displayed;
CC Name=2;
CC IsoId=P20701-2; Sequence=VSP_002738;
CC Note=No experimental confirmation available;
CC -!- TISSUE SPECIFICITY: Leukocytes.
CC -!- DOMAIN: The integrin I-domain (insert) is a VWFA domain. Integrins
CC with I-domains do not undergo protease cleavage.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- DATABASE: NAME=PROV; NOTE=CD guide CD11a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd11a.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by, and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; Y00796; CAA68747.1; --  
 EMBL; AC002310; AAC31672.1; --  
 PIR; S03308; S03308.  
 PDB; 1CQP; X-ray; A/B=153-334.  
 PDB; 1DQG; NMR; A=149-336.  
 PDB; 1LFA; X-ray; A/B=--  
 PDB; 1MJN; X-ray; A=153-331.  
 PDB; 1MQ8; X-ray; B/D=155-331.  
 PDB; 1MQ9; X-ray; A=152-330.  
 PDB; 1MQA; X-ray; A=152-330.  
 PDB; 1ZON; X-ray; @=150-336.  
 PDB; 1ZOO; X-ray; A/B=150-336.  
 PDB; 1ZOP; X-ray; A/B=150-336.  
 Genew; HGNC:6148; ITGAL.  
 MIM; 153370; --  
 GO; GO:008305; C:integrin complex; TAS.  
 GO; GO:006928; P:cell motility; TAS.  
 InterPro; IPR00413; Integrin\_alpha.  
 InterPro; IPR002035; VWF\_A.  
 Pfam; PF01839; FG-GAP; 3.  
 Pfam; PF00357; Integrin\_alpha; 1.  
 Pfam; PF00092; VWA; 1.  
 PRINTS; PR01185; INTEGRINA.  
 PRINTS; PR00453; VWFADOMAIN.  
 SMART; SM00191; Int\_alpha; 5.  
 SMART; SM00327; VWA; 1.  
 PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 PROSITE; PS0234; VWFA; 1.  
 3D-structure; Alternative splicing; Calcium; Cell adhesion;  
 Direct protein sequencing; Glycoprotein; Integrin; Magnesium;  
 Receptor; Repeat; Signal; Transmembrane.  
 SIGNAL 1 25  
 CHAIN 26 1170 Integrin alpha-L.  
 DOMAIN 26 1090 Extracellular (Potential).  
 TRANSEM 1091 1110 Potential.  
 DOMAIN 1112 1170 Cytoplasmic (Potential).  
 REPEAT 42 91 FG-GAP 1.  
 REPEAT 92 149 FG-GAP 2.  
 DOMAIN 156 327 VWFA.  
 REPEAT 350 400 FG-GAP 3.  
 REPEAT 401 455 FG-GAP 4.  
 REPEAT 457 516 FG-GAP 5.  
 REPEAT 518 575 FG-GAP 6.  
 REPEAT 578 630 FG-GAP 7.  
 CA\_BIND 468 476 Potential.  
 CA\_BIND 530 538 Potential.  
 CA\_BIND 590 598 Potential.  
 SITE 1115 1119 GFFKR motif.  
 DISULFID 73 80 By similarity.  
 DISULFID 111 129 By similarity.  
 DISULFID 653 707 By similarity.  
 DISULFID 771 777 By similarity.  
 DISULFID 845 861 By similarity.  
 DISULFID 998 1013 By similarity.  
 DISULFID 1021 1052 By similarity.  
 CARBOHYD 65 65 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 89 89 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 188 188 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 649 649 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 670 670 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 726 726 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 730 730 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 862 862 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 885 885 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 897 897 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 1060 1060 N-linked (GlcNAc... ) (Potential).  
 CARBOHYD 1071 1071 N-linked (GlcNAc... ) (Potential).

Q -> QGVHGLVEMOTSKOILCRPAGDAEHTVGAQEGELPC  
 PMGYSEAFRDNIRAGPCR (in isoform 2).  
 R -> W (in Ref. 1 and 2).  
 Y -> I (in Ref. 2).

FT VARSPLIC 954 954  
 FT 214 214  
 FT CONFLICT 660 660  
 FT CONFLICT 155 155  
 FT STRAND 164 164  
 FT TURN 165 166  
 FT HELIX 169 185  
 FT TURN 186 188  
 FT STRAND 191 198  
 FT STRAND 202 206  
 FT STRAND 208 214  
 FT HELIX 217 221  
 FT TURN 222 223  
 FT STRAND 229 229  
 FT HELIX 233 243  
 FT TURN 244 244  
 FT HELIX 247 249  
 FT TURN 250 250  
 FT TURN 253 254  
 FT STRAND 256 263  
 FT HELIX 274 276  
 FT TURN 277 278  
 FT STRAND 280 286  
 FT HELIX 288 290  
 FT HELIX 293 297  
 FT TURN 298 299  
 FT HELIX 300 302  
 FT HELIX 307 310  
 FT STRAND 311 314  
 FT HELIX 317 319  
 FT HELIX 323 330  
 SQ SEQUENCE 1170 AA; 128819 MW; 39A7AF92EF286FC0 CRC64;

Query Match 26.4%; Score 1552.5; DB 1; Length 1170;  
 Best Local Similarity 34.5%; Pred. No. 1.4e-93;  
 Matches 405; Conservative 208; Mismatches 460; Indels 101; Gaps 37;

Qy 1 FNLDTENAMTFQ--ENARGFGQSVVQLQGSRVVVGAPQEIIVANQORSLVQCDYSTGSCE 58  
 Db 26 YNLDRVARGSFSPPRAGRHFGYRVLQV-GNGVIVGAPGE---GNSTGSLYQCQSGTGHCL 81  
 Qy 59 PIRLOVPVEAVNMSLGLSLAATTSPOLLACGPTVHOTCSNTVVKGLCFLFGNLR--- 115  
 Db 82 PVLUR-GSNTYSKYLGMILATDPTDGSILACDGLSRTCDQNTYLSGLCYLFRNLQGP 140  
 Qy 116 -QQQKPFPEALRGCPQSDIAFLIDGSGSIIIPHDFFRMKEFVSTVMEQKLKSTLFSLM 174  
 Db 141 LQGRPGFQECIKG---NVDLVFLDGSMSLQDPDEFQKILDFFMKDVMKKLSNTSYQPA 196  
 Qy 175 QYSEEPRIHFTKFEFQNNPNRSLVKPITOLLGRTHATGVKRVIRELLNITNGARKNAF 234  
 Db 197 QFSTSYKTEFDSDYVRKQDPDALLKHVKHMLLLTNTFGAINVYVATEVFREELGAPD 256  
 Qy 235 KILIVITDGKFGDPLGYEVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPERD 294  
 Db 257 KVLIIITDGE--ATDSNIDAKD-----IIRYIIGIKGFQTKESQETLHKFASKPASE 309  
 Qy 295 HVFQVNNFEALKTIQNOLREKIFAIEGTQTGSSSFEHEMSQEGFSAAITSNGLLSTVG 354  
 Db 310 FVKILDTFEXLKDILFTELQKKIYVIEGTSKQDLTSFNKMLSSSGISADLSRGHAVVG 369  
 Qy 355 SYDWAGVVF-LYTSKEXSTINMTRVDSMDNDAYLGAAA-IILNRVQSVILGAPRYQH 412  
 Db 370 AKDWAGGFLDLKADLQDDTFIGNEPLTPEVRAGYLGTYVTWLPFSRQKTSLLASGAP 429  
 Qy 413 IGLVAMPR--QNTGMWESNANVKGTQIGAYFGASLCSDVDVDSNSTDLVLIGAPHYEOT 470  
 Db 430 MGRVLLFQEPQGGHNSQVQTIHGTQLGSYFGGELCGVDVDQDGETELLIGALFLYGEQ 489  
 Qy 471 RGGQSVVCPPLPRGQARWQCDAV--LYGEOQGPWGRFGAALTVLGDVNGDKLTDVAIGAP 528

Db 490 RGRVFYIY-----QRRQLGFEVSELSQDGPYPLGRFGEAITALTIDINGDLVDVAVGAP 544  
 QY 529 GEDNKGAVLFGHTSGSISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTV 588  
 Db 545 LEE--QCAVYFNGRG-GLSPQSQRIEGTQVLSGIQWFGSIHGKVLGEGGLADVAV 601  
 QY 589 GAGCHVLLRSQSVLRVKAIMEPNPREVARNVPECNDQVV-KGKEAGEVRVCLHVOKSTR 647  
 Db 602 GABSONVILSSRPVDMVITLMSFSPAEIPVHEVECSYSTSNKMKEGVNITICFQI-KSLY 660  
 QY 648 DLREGIOQSVVYDIALDSGRPHSRVAFNETKNSTRQTVLGLTQTCTELKLQLPNCI 707  
 Db 661 PQ-QCRLVANLYTYLQDGHRTRRRGLPGGRHRLRNIAVT-TSMSCDTDFSFHFCV 718  
 QY 708 EDVPSVILVLFNSL---VCTPLS--AFGN-----LRPVLAEDAQRILFTALFFPEKNCGN 757  
 Db 719 QDLISPINVLSNLSLWEEGTPTDQAQGDIPILRPSLHSETWBI-----PFEKNCGE 773  
 QY 758 DNICQDLSITTFMSLDCLVGGPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKVS 817  
 Db 774 DKCEANLRVSFSPARSRLRLTAPASLSVELSLNLEEDAYVVDLHPPGLSRKVE 833  
 QY 818 TLQORSORSRLACRS--ASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKAS 875  
 Db 834 ML---KPHSQIPVSCBELPEESRLGRAL---SCNVSSPIFKAGHSVALQMMFNTLVNS 887  
 QY 876 LGNKLKLLKANVTSENN---MPTNKTFFOLELPVKYAVVMVTSHGVS TKYLNFTASEN 931  
 Db 888 WGSDELHANVTCCNEDSLLDENSATTI---IPILYPINILQDOEDSTLYVSFTPKGP 944  
 QY 932 TSRVMQHOYQV---SNLQORSLP-ISLVFLVPVRLNQTIVDRPQVTFSENLSSTCHTK- 986  
 Db 945 KIHQVKHMVQVRTOPIHHDNIETLEAVGVQPPSEGPITHQWSVQMEPPV--PCHYED 1002  
 QY 987 -ERLPSHD--FLAELRKAPVNVNCISAVCORIQDIPFGIOBEFNATLKNLSFDWYIK 1043  
 Db 1003 LERLPDAAEPCLPALFRCPW-----FRQELVQVIGTLELVEIE 1044  
 QY 1044 TSHNLLIYSTAILFNDVSFTLLPGCGAFVRQTEKVPFEPVPLPLIVGSSVGGLL 1103  
 Db 1045 AS-SMFSLSLSLSISNSSKHFLYGSNASL-AQVVMKVDVYVEKQMLYLYLVLSGIGLL 1102  
 QY 1104 LLALITAAALYKLGFFKRYQKDMMESEG-GPPGAP 1136  
 Db 1103 LLLLLLIVLYKGVFFKRLKMEAGRGVNGIP 1136  
 RESULT 13  
 Q6TVB8  
 ID Q6TVB8 PRELIMINARY; PRT; 1166 AA.  
 AC Q6TVB8  
 DT 05-JUL-2004 (TREMblrel. 27, Created)  
 DT 05-JUL-2004 (TREMblrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMblrel. 27, Last annotation update)  
 DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N A.  
 RA Dileepan T. Thumkikat P., Kannan M.S., Maheswaran S.K.;  
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (by similarity).  
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.  
 DR EMBL; AY382558; AAQ90015.1; -  
 DR InterPro; IPR004113; Integrin\_alpha.  
 DR InterPro; IPR002035; VWF\_A.  
 DR Pfam; PF01839; FG-GAP; 1.  
 DR Pfam; PF00357; Integrin\_alpha; 1.  
 DR Pfam; PF00092; VWA; 1.  
 DR PRINTS; PR01185; INTEGRINA.

DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00191; Int\_alpha; 5.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00242; INTEGRIN\_ALPHA; 1.  
 DR PROSITE; PS00234; VWEA; 1.  
 KW Cell adhesion; Integrin; Transmembrane.  
 SQ SSEQUENCE 1166 AA; 128723 MW; 7B8D8AFBA896C9DF CRC64;  
 Query Match 26.4%; Score 1552; DB 2; Length 1166;  
 Best local similarity 34.4%; Pred. No. 1.5e-93;  
 Matches 407; Conservative 200; Mismatches 469; Indels 106; Gaps 33;  
 QY 1 ENLDTENAMTFOENARG--FGQSVVOLQSGRVVVGAPQEIIVAAQSGSLYQCDYSTGSC 58  
 Db 25 YNLDVRHVQNFSPPLAGRHFGVRLQV-GNVVVVGAPSE---GNSMGNLYQCOPEFGDCL 80  
 QY 59 PIRLOVPVBAVNMSLGLSLAAATTPPQLLACGPTVHQTCSENTYVRKGLCFGLFSGNL 118  
 Db 81 PVTLS--SNVTSKYLQWTLATDPTSDNLLACDPGLSRCTCDQNIYLSGLCYLIHENLRGP 138  
 QY 119 QKPEALRGCPQEDSIAFLIDGSGSIIIPDPRMKEFVSTVMEQLKSKTILFSLMOYSE 178  
 Db 139 LQHPGYQCIKGNVDLVFLFDGMSLQDDEFEKIVDFMKDVMKLSNSYQFAAQFST 198  
 QY 179 EPRIHETFEFQNNPRLVKPITOLLGRTHATGVRKVIRELLNITNGARKNAFKILI 238  
 Db 199 YFTEFTFDYIKQDPDALLAGVKHMLLTNTFGAINYVAKVFRPDLGARDAKYVLI 258  
 QY 239 VITDGKFGDPLGYEDVIEPADREGVIRVYGVGDAPRSEKSEKSOELNTIASKPRDRHVQ 298  
 Db 259 IITDGK--PPTNTTLMRPKTSS---RSLIGIGKFKTKESQEALHQFASKPVEBFVKI 311  
 QY 299 VNNFEALKTIONLRKIFAIEGTQTGSSSSPHEMSQEGSAITNSGNLLSTVGSYDW 358  
 Db 312 LDTFEKLKDLFTLOKKIYVIEGTQDLTSFMELSSSGISADLSEGHGAVGAGAKDW 371  
 QY 359 AGGVF-LYTSKEKSTFINTRVDSMDNDAVLGYAAA-IILNRVQSLVLCAPRYQHIGLV 416  
 Db 372 AGFLDLKADLKSSTFVNEPLTVESRAGLYGVTVTRLPSTRGTMSLLATGAPKQHVGRV 431  
 QY 417 AMFRQ--NTGMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVLIGAPHYYEOTRGQ 474  
 Db 432 LLFQPKRGKPSQIQEIDGIQIGSYFGELCGVDVDRDGETELLIAAPLYYGEQGRG 491  
 QY 475 VSVCPILPRGQARWQCDVLYGEOGPGRFGNALTVLGDVNGDKLTDVAIGAPGEDNR 534  
 Db 492 VFYIY---QKIQLFQVMSLQSETGYPLGRFGAAIAALTIDINGDELTDVAVAPLEE--Q 546  
 QY 535 GAVYLFHGTSGSGISPSHSORIASKLSPLQYFGQSLSGQDLTMDGLVDLTVGAGQHV 594  
 Db 547 GAVYIFNGQOG-GLSPRPSQRIEGTQMFSGIQWFGSIHGKVLGEGGLADVAVAGQV 605  
 QY 595 LLRSQVILRVKAIMEFPNPREVARNVPECNDQVVKGEAG-EVVRVCLHVQK--STRDLR 651  
 Db 606 IVLSSRPVVDIITSVSFPAETPVHEVECSYSTSNQKEGVNLTVCVQFKVSLIST---F 661  
 QY 652 EGOIQSVVYDIALDSGRPHSRVAFNETKNSTRQTVLGLTQTCTELKLQLENCIEDPV 711  
 Db 662 QGHVLANLYTYLQDGHRTSRGLFFGGRHKLIGNTAVTPV-KSCFVFMVFFHFCIQLLI 720  
 QY 712 SPITVRLNFSL---VGTPLS--AFGNLRPVLAEDAQRILFTALPPFEKNCNDNICODDLS 766  
 Db 721 SPINVSLSLWEEGTPTDPRALDRDIIPILKPSHLETKEIPFENCGEDKNCXADLK 780  
 QY 767 ITFSFMSLDCLVVGGRPREFNVTVTRNDGEDSYRTQVTFPPPLDLSYRKYSTLQNSQR 826  
 Db 781 LAFSDMRSKILRLTPSASLSVRLTLTAEDAAYVQVTLSPQGLSFRKVEIL--KPHS 837  
 QY 827 SWRLACESASSTEVSGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKANV 886  
 Db 838 HVPVGCCELPEEAHVHS-RALSCNVSPFSGDSMVDIQVMFNTLQKSGWGDPIELQANV 896  
 QY 887 TS-----ENNMPRTNKTFFOLELPVKYAVVMVTSHGVS TKYLNFTASENTRVMQH 938

Db 897 SCNNEDSLLEDNGSATTSS-----IPVMYPIVLTQKQENSTLYISFTPKSPRIHHVKH 949  
QY 939 QYQV-----SNLQGRSLPISLVLPVRLNQTVI---WD---RPOVTFSS-ENLSSTCHTKE 987  
Db 950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESFSDAE 1008  
QY 988 RLPSSHDFLAELRKAPVNCNSIAVCQRIQCDIPFFGQIEEPNATLKNLSFDWIKTSHN 1047  
Db 1009 -----SCSFGT--EPRCPIDF---RQELVQVNGMVELRGITIKAS-S 1044  
QY 1048 HLLIVSTAELFNDSVFTLLPGOGAFVRSQTETKVEPFEVNPPLIIVGSSVGGILLIAL 1107  
Db 1045 MSLCSSLAISFNSSKHFHLHGRNASM-AQVVMKVDLYVYKEMLYLVLSIGIGILLFL 1103  
QY 1108 ITAALYKLGPFKQYKDM-----SEGGPPGAEPPQ 1137  
Db 1104 IFIALLYKVGFFKRLKEKMEANVDASSEIPGEDAGQPELEKE 1145

## RESULT 14

AAQ90015

ID AAQ90015 PRELIMINARY; PRT; 1166 AA.

AC AAQ90015;

DT 02-MAR-2004 (TrEMBLrel. 27, Created)

DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)

DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)

DE Lymphocyte function-associated antigen 1 alpha subunit CD11a.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RA Dileepan T., Thumbikat P., Kannan M.S., Maheswaran S.K.;

RT "Molecular cloning and sequencing of bovine CD11a.";

RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY382558; AAQ90015.1; -

SQ SEQUENCE 1166 AA; 128723 MW; 7B8D8AFA896C9DF CRC64;

## Query Match

Best Local Similarity 34.4%; Pred. No. 1.5e-93;

Matches 407; Conservative 200; Mismatches 469; Indels 106; Gaps 33;

QY 1 FNLDENAMTFQENARG--FCQSVVLOGSVVVGAPQEIIVAAQNRGSLYQCDYSTGCE 58

Db 25 YNLDRHVQNPFPFLAGRHFGYRVLQV-GNGVVVGAPSE---GNSMGNLYQCQPETGDL 80

QY 59 PIRLQVPVEAVNMSIGLSLAATTSPOLLACGPTVHOTCSENTYVKGLCFLFGSNLRQP 118

Db 81 PVTLSS--SNYTSKYLGMWLTADPTSDNLLACDGLSRTCDQNIYLSGLCYLIHENLRGPV 138

QY 119 QKPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMQKKSKTLFLSLMYS 178

Db 139 LQGHFGYQECIKGNVDVFLFDGMSLQQDEFEKIVDFMKDVKMKLSNYSYQFAAVQFST 198

QY 179 EFRTHFTKEPQNNPNSRLVKPITQLLGRHTATGVRKVIKRELLNITNGARKNAFKILI 238

Db 199 YFRTEFTLDYIKOKDPDALLAGVHKMLLNTFGAINYAKVFRPDLGLARPATKVLII 258

QY 239 VITDGEKFGDPLGYEDVIEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPDRHVQ 298

Db 259 IITDGG--PPTNTLMRPKTS---RSLIGIGKFKTESQEAHLQFASKPVEEFVKI 311

QY 299 VNNFEALKTQNRQREKIFALETGTGSSSSFEHMSQEGPSAINTGNPLLSYGVYDW 358

Db 312 LDTPEKLDKLTQELQKIIYVEGTSKQDLTSTFNMLSSSGISADLSEGHGVVGVAGAKDW 371

QY 359 AGGVF-LYTSKEKSTFTNTMTVDSDMDNDAYLGAAA-IILENNRQSVILGAPRYOHLGLV 416

Db 372 AGGFLDLKADLKSFTFVNEPLTVESRAGYLGTVTLRPLSRGTWSLLATGAPKYQHVGRV 431

QY 417 AMFRQ--NTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYQTRGGQ 474  
Db 432 LIFQOPKRGFPWQIOEIDGIGQISYFGGELCGVDVDRDGETELLTAAPLYYGEQGR 491  
QY 475 VSVCLPFRGQARWQCDALVYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNR 534  
Db 492 VFYI---QKIQLBQFQWSELQGETGYPLGRFGAAIALTDINGDELTDVAVGAPLEE--Q 546  
QY 535 GAVYLFHTSGSGISPHSQRISAGSKLSPLOYFGOSLSGGQDLMTGLVDLTVGAQOHHV 594  
Db 547 GAVYIFNGQQG-GLSPRPSQRIEQTQMPFGIOWFGRSIHGKDLGGDLGADVAVGAEQV 605  
QY 595 LLRSQVLRVKAIMEENPREVARNPFCNDQVVVKKEAG-EVRVCLHVQK--STRDLRL 651  
Db 606 IYLSRPVVDIITSVSFPAEIPVHEVECSYSTNQKKEGVNLIVCQVKSLIST---F 661  
QY 652 EGQIQSVVYTDALDSGRPHSRVAFNETKNSRSTQTVLGTQTCETLKLQLPNCIEDPV 711  
Db 662 QGHLVANLTYTLQDGHRTSRGLFPGGKHKLIGNTAVTPV-KSCFVFWFHPFICIQDLI 720  
QY 712 SPIVLRNFSL---VGTPLS--AFGNLRPVLAEADAQRLFTALFPEKNCNDNICODDLS 766  
Db 721 SPINVSUSYSLWBEETPRDRALDRIPPLIKPSPHLETKEIPFNKNCGEDKNCEADLK 780  
QY 767 ITFSFMSLDCLVVGGPREFNVTVTRNDGDSYRTQVTFPPPLDLSYRKVSTLQNRQSR 826  
Db 781 LAFSDMSKILRLTPSASLSVLRLTNTAEDAYWQVTLSPFQGLSPRKVEIL---KPHS 837  
QY 827 SWRIACSASSTEVSGALKSTCSINHPIPEENSEVTNITFDVDSKASLGNKLLKANV 886  
Db 838 HVPVGCBELEPEEAVVHS-RALSCNVSPFIFGDSMVDIQVMFNTLQKSGWGDFTLOANV 896  
QY 887 TS-----ENNMPRTNKTEFQLELPVKYAVVMVTVSHGVSTKYLNTASENTRVQMH 938  
Db 897 SCNNEDSLLEDNSATTS-----IPVMYPIVLTQKQENSTLYISFTPKSPRIHHVKH 949  
QY 939 QYQV-----SNLQGRSLPISLVLPVRLNQTVI---WD---RPOVTFSS-ENLSSTCHTKE 987  
Db 950 IYQVRIQPSNYDNMP-PLEALVRVPRVHSEGLITHKWSIQMEPPVNCSPRNLESFSDAE 1008  
QY 988 RLPSSHDFLAELRKAPVNCNSIAVCQRIQCDIPFFGQIEEPNATLKNLSFDWIKTSHN 1047  
Db 1009 -----SCSFGT--EPRCPIDF---RQELVQVNGMVELRGITIKAS-S 1044  
QY 1048 HLLIVSTAELFNDSVFTLLPGOGAFVRSQTETKVEPFEVNPPLIIVGSSVGGILLIAL 1107  
Db 1045 MSLCSSLAISFNSSKHFHLHGRNASM-AQVVMKVDLYVYKEMLYLVLSIGIGILLFL 1103  
QY 1108 ITAALYKLGPFKQYKDM-----SEGGPPGAEPPQ 1137  
Db 1104 IFIALLYKVGFFKRLKEKMEANVDASSEIPGEDAGQPELEKE 1145

## RESULT 15

ITAL MOUSE

ID ITAL MOUSE

AC P24063; STANDARD; PRT; 1163 AA.

DT 01-MAR-1992 (Rel. 21, Created)

DT 01-MAR-1992 (Rel. 21, Last sequence update)

DT 05-JUL-2004 (Rel. 44, Last annotation update)

DE Integrin alpha-L precursor (Leukocyte adhesion glycoprotein LFA-1

DE alpha chain) (LFA-1A) (Leukocyte function associated molecule 1, alpha

DE chain) (CD11a).

GN Name=Itgal; Synonyms=Lfa-1;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91268576; PubMed=2051027;

RA Kaufmann Y., Tseng E., Springer T.A.;

RT "Cloning of the murine lymphocyte function-associated molecule-1

FT	DISULFID	840	856	By similarity.
FT	DISULFID	993	1009	By similarity.
FT	DISULFID	1017	1048	By similarity.
FT	CARBOHYD	86	86	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	185	185	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	270	270	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	444	444	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	668	668	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	696	696	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	724	724	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	728	728	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	776	776	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	857	857	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	880	880	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	890	890	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	899	899	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	927	927	N-linked (GlcNAc. . .) (Potential).
FT	CARBOHYD	1056	1056	N-linked (GlcNAc. . .) (Potential).
SEQ	SEQUENCE	1163 AA; 128343 MW; A7A3078489E8232F CRC64;		
Query Match				
Best Local Similarity 26.1%; Score 1533.5; DB 1; Length 1163;				
Matches 398; Conservative 217; Mismatches 458; Indels 99; Gaps 36;				
Qy	1	FNDTENAMTFQENA-RGFQGSVVLQGSRRVVVGAPOEIVAAANQGRSLYQCDYSTGSC	59	
Db	24	YLNDRTPQSLAQRHFGVQLIEDG-VVVGAPGE---GDNTGGLYHCRSTSSEFCQ	79	
Qy	60	IRLQVPVEAVNMSLGLSLAATTSPQALLACGPTVHQTCSNTYVKGCLPFGSNLRQOQ	119	
Db	80	VSLH-GSNHTSKYLGMTLATDAKGSLLACDPLGSRTCQNTYLSGLCYLPQSLG	138	
Qy	120	KFPEALRGCPQEDSDIAFLIDGSGSIIPHPFRMKPEFVSTWMEOLKSKTLFSLMOYSEE	179	
Db	139	QNRPAYECMKGVDLVFLFDGSGSLDRKDFEKLIFPMKDMVKLSNTSTQFAAVQSF	198	
Qy	180	FRHFTPEKF-ONNPNSRLVKPITOLLGRTHRTATGVRKVIKIRELLINTNGARKNAFKILI	238	
Db	199	CRTEFTFLDYKQNKNDVLLGSVQPMFLTNTFRAINVYVAHVFEESGARPDKVIV	258	
Qy	239	VITDGEKFGDPLGVEDVIPADREG-----VIRYVIGVGDAFRSKSKSROELMTIASK	291	
Db	259	IITDG-----EASDKGNISAAHDIPTYIIGIKHFVSVQKQKTLHFASEP	304	
Qy	292	PROHVFQVNFEEALKTIQNOLREKIFAIEGTQCTSSSEHMSQEGFSAITSNGLPLS	351	
Db	305	VEEFVKILDTFEKUKDUFTLQRRYIAIEGTNRDLTSFNNELSSSGISADLSKGHVA	364	
Qy	352	TVGSYDMWAGGVF-LYTSKEKSTFTNTRVDSMDMDAYLGAAA-IILNRVQSLVLGAPR	409	
Db	365	AVGAKDWAGGFLDLREDLQGNATFVGQPLTSDVEGGVLYGVYVAMWTSRSPPLAAGAPR	424	
Qy	410	YOHIGLVAMPR--ONTQWMSNANVKQTQIYGFGASLCSVDVDSNGSTDLVLGAPHYY	467	
Db	425	YQHVGVQLLFPQPEAGGRWNQTKIEGTQICSYFGGELCSVDLDQDGEAEILLIGAPLFF	484	
Qy	468	EQTEGGGVSCPLPRGQARWQCDVILYGGQGPWGFAGALTVLGVNGDKLTDVAIGA	527	
Db	485	GEORGGRVFTY--QRQSGFEMVSELOQDPGVPLGFRGAAITALTIDINGRLTIDVAVGA	541	
Qy	528	PGEDNRGAVYLFHGTSGSITSPSHSQRIAGSKLSPLRQVFGQSLGGQDLTMQGLVDLT	587	
Db	542	PLEE--QGAVYIENGKPG-GLSQPSQRIQGAQVFGIRWFGRSIHGVKDLGGDLADVV	598	
Qy	588	VGAQGHVLLRSQPVLRVKAIMEFNPREVARNVFECNDQVVKGEAG-EVRVCIHFVKST	646	
Db	599	VGAEGRVVLLSRPVDVVTLSFSPPEIIPVHEVECSYSAREEQKHGVKLCACFRKPLT	658	
Qy	647	RDLRLEQIOQSVTYDLALDSGRPHSRAVFNETHKNSTRQTVGLGLTQTCETLKLQIPNC	706	
Db	659	PQ--FQGRLLANISYTLQLDGHRMRSRGLFPDGSHELMSGNTSITTP-DKSLDFHFFPIC	715	
Qy	707	IEDPVSPIVLRNLFSIV--GTPLSARGN-LRPVIAEDAORIETALPPEFKNCNDNICO	762	



Db 716 IQDLISPINVSLNLSLEEETPRQGRAMQPIILRPSIHV-TKEIPFEKNCGEDKKCE 774  
Qy 763 DDLISITFSFMDLCLVVGPP-----REFNVTIVRNDGEDSYRTQVTFPPPLDLSYRKV 816  
Db 775 ANLTISSPARS-----GPLRLMSSASLAVEWTLNSGEDAYWVRLDLDPFRLSPRKV 827  
Qy 817 STLQNRQSRWRLACESASSTEVS GAL-KTSCSINHPIPPENSEVTFNITFDVDSKAS 875  
Db 828 EMLQ---PHSRMPVSCBEL--TEGSSLLTKILKCNVSSPIFKAGQEVSLQWFMFTLNSS 882  
Qy 876 LGNKLLKANVTSEN-NMPTNKTEFQLELPKYAVVMVVTSHGVSTKYLNFTASENISR 934  
Db 883 WEDEVELNGTVHCENENSSLOEDNSAATHIPVLYPVNILTKEQENSTLIYISFTPKGPKTQ 942  
Qy 935 VMQHQYQVSNLQORSPLSILVLPVRLNQTVIWRDPO----VTFSENLS-----TCHTK 986  
Db 943 QVOHVYQV-----RIQPSAYDHNMT-LEALVGVPFRHSEDLITYTWSVQTDPLVTCHSE 996  
Qy 987 E-RLPSHSDFLAELRKAPVNVNCIAVCORICDIPFGIOEEFNATLKGNLSEFDWYIKTS 1045  
Db 997 LKRPSE---ABQPCLEGV-----QFRCPVFI---RWEILIOVTGTVELSKEIKAS 1042  
Qy 1046 HNHLIVSTABILFNDVSFTLLPGQGAFFVRSQTEKVEPFPVNPPLPLIVGSSVGGLLLL 1105  
Db 1043 -STLSLCSLSVSFNSKHFHYGSKA-SEAQLVKVDLIHEKEMLVVYVLSGIGGLVLL 1100  
Qy 1106 ALITAAALYKLGFFKQYKDM-SEGPPGAEP 1136  
Db 1101 FLIFIALYKVGFFKRNLERMEADGGVPNGSP 1132

Search completed: November 9, 2004, 12:17:43  
Job time : 141.75 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:45:42 ; Search time 103.25 Seconds  
(without alignments)  
3950.365 Million cell updates/sec

Title: US-09-902-481b-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGGPGAEPEQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5868	99.9	1153	2	AAW65090 Human Bet
2	5868	99.9	1153	3	AAW65090 Human Bet
3	5868	99.9	1153	5	AAU0252 Human int
4	5868	99.9	1153	5	ABG61469 Human Bet
5	5868	99.9	1153	7	AAO14428 Integrin
6	5868	99.9	1153	7	ADD25615 Binding d
7	5858	99.7	1153	2	AAW65091 Human Bet
8	5852.5	99.6	1152	8	ADP12435 Protein e
9	5845.5	99.5	1152	8	ADP12435 Protein e
10	3487	59.4	1163	8	ADP44061 Human CDI
11	3473	59.1	1163	8	ADN02004 Human inf
12	3473	59.1	1163	8	ADQ17510 Human sof
13	3464	59.0	1163	2	AAW65091 Human Bet
14	3450	58.7	1163	2	AAW65091 Human Bet
15	3450	58.7	1163	3	AAW65091 Human Bet
16	3450	58.7	1163	5	ABG61470 Human Bet
17	3448	58.7	1163	6	ABU07406 Protein d
18	3448	58.7	1163	7	ADG32005 Human hom
19	3417	58.2	1161	2	AAW65091 Human Bet
20	3417	58.2	1161	2	AAW65091 Human Bet
21	3417	58.2	1161	2	AAW65091 Human Bet
22	3417	58.2	1161	2	AAW65091 Human Bet
23	3417	58.2	1161	2	AAW65091 Human Bet
24	3417	58.2	1161	2	AAW65091 Human Bet
25	3417	58.2	1161	3	AAW65091 Human Bet

## ALIGNMENTS

RESULT 1  
AAW65090  
ID AAW65090 standard; protein; 1153 AA.  
XX  
AC AAW65090;  
XX  
DT 28-SEP-1998 (first entry)  
XX  
DE Human Beta-integrin CD11b subunit protein.  
XX  
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
KW lung inflammation; acute respiratory distress syndrome; CD11b subunit;  
KW rheumatoid arthritis.  
XX  
OS Homo sapiens.  
XX  
PN US5728533-A.  
XX  
PD 17-MAR-1998.  
XX  
PF 07-JUN-1995; 95US-00485618.  
XX  
PR 23-DEC-1993; 93US-00173497.  
PR 05-AUG-1994; 94US-00286889.  
PR 21-DEC-1994; 94US-00362652.  
XX  
PA (ICOS-) ICOS CORP.  
XX  
PI Van Der Vieren M, Gallatin WM;  
XX  
DR WPI; 1998-206565/18.  
XX  
PT Screening assay for modulators of integrin binding - using immobilised or  
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.  
XX  
PS Example 5; Fig 1A-D; 106pp; English.  
XX  
CC This sequence represents a human beta-integrin CD11b subunit which is  
CC used to describe a method for identifying compounds that modulate the  
CC interaction of the beta-integrin alpha-d subunit with a binding partner  
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
CC -d binding partner, one of which is immobilised and the other of which is  
CC labelled, in the presence of a test compound, and determining if the  
CC compound affects binding between the alpha-d polypeptide and alpha-d  
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
CC comprising the cytoplasmic, transmembrane or extracellular domain of  
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat

Abg61468 Human Bet  
Aaw23064 Human bet  
Aaw65106 Human Bet  
Aaw72837 Human alp  
Aaw73343 Human alp  
Aab07376 Human alp  
Abg61485 Human Bet  
Aar78169 Rat alpha  
Aaw23062 Rat beta  
Aaw60004 Rat alpha  
Aaw72824 Rat alpha  
Aab07374 Rat Beta2  
Abg61483 Rat Beta2  
Aaw65104 Rat beta-  
Aaw73345 Rat alpha  
Aaw23061 Mouse bet  
Aaw60003 Mouse alp  
Aaw65103 Mouse bet  
Aaw72836 Mouse alp  
Aaw73347 Mouse alp

CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
XX and rheumatoid arthritis

SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 2; Length 1153;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQIIVAAQNGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFQENARGFQSVVQLQGSRVVVGAPQIIVAAQNGSLYQCDYSTGSCPEI 76

QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCTSENYYVKGFLFGSNLRQOPQK 120  
DB 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCTSENYYVKGFLFGSNLRQOPQK 136

QY 121 FPEALRCGPQEDSDIAFLIDGSGSIIIPDFFRMKEFVSTWQELKKSKTLFSLMOYSEEF 180  
DB 137 FPEALRCGPQEDSDIAFLIDGSGSIIIPDFFRMKEFVSTWQELKKSKTLFSLMOYSEEF 196

QY 181 RIHFTPFQNNPRSLVKPIITQLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI 240  
DB 197 RIHFTPFQNNPRSLVKPIITQLLGRTHATGIRKVVRELFNITNGARKNAFKILIVI 256

QY 241 TDGEKFGDPLIGYEDVIEADREGVIRVIVGVDAPRSEKROELNITASKPRDRHVPQVN 300  
DB 257 TDGEKFGDPLIGYEDVIEADREGVIRVIVGVDAPRSEKROELNITASKPRDRHVPQVN 316

QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQEGFSAAITNSGFLPSTVGSYDWAG 360  
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSPHEHMSQEGFSAAITNSGFLPSTVGSYDWAG 376

QY 361 GVFLYTSKSKSTFTINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYOHIGLVAMFR 420  
DB 377 GVFLYTSKSKSTFTINMTRVDSMDNDAYLGAAAILLRNVQSLVGLGAPRYOHIGLVAMFR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVILIGAPHYYEOTRGQSVCP 480  
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVILIGAPHYYEOTRGQSVCP 496

QY 481 PRGORARWQCDVLYGQGPQWGRFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVLYF 540  
DB 497 PRGORARWQCDVLYGQGPQWGRFGAALTVLGVNGDKLTDVAIGAPGEDNKGAVLYF 556

QY 541 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGGAQGHVLLRSQ 600  
DB 557 HGTSGSGISPSHQSRIAGSKLSPLQYFGQSLGGQDLTMDGLVDLTGGAQGHVLLRSQ 616

QY 601 PVLRVKALMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTDRLRREGIQSVWT 660  
DB 617 PVLRVKALMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTDRLRREGIQSVWT 676

QY 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLENCIEDPVSPTVLRINF 720  
DB 677 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLENCIEDPVSPTVLRINF 736

QY 721 SLVGTPLSAGNLRPVLAEDAQRFTALFPPEKNCNDNICQDDLSITTFMSLDCLVWG 780  
DB 737 SLVGTPLSAGNLRPVLAEDAQRFTALFPPEKNCNDNICQDDLSITTFMSLDCLVWG 796

QY 781 GPREFNVTVRNDGDSYSTQVTFEPFLDLSYKVKSTLQNSORSRWRACESASSTEV 840  
DB 797 GPREFNVTVRNDGDSYSTQVTFEPFLDLSYKVKSTLQNSORSRWRACESASSTEV 856

QY 841 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 900  
DB 857 SGALKSTSCSINHPIPFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPRTNKTET 916

QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSPLISLVFLVPV 960  
DB 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMOHQVQVSNLQORSPLISLVFLVPV 976

QY 961 RLNQTVIWDPRQVTFSENLSTCHTKERLPBSHSDFLAELRKAPVWVNCISIAVCORIQCDIP 1020  
DB 977 RLNQTVIWDPRQVTFSENLSTCHTKERLPBSHSDFLAELRKAPVWVNCISIAVCORIQCDIP 1036  
QY 1021 FFGIQBEFNATLKGNLISFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTET 1080  
DB 1037 FFGIQBEFNATLKGNLISFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQCAFVRSQTET 1096  
QY 1081 KVPEFVEPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMEGPPGAEPPQ 1137  
DB 1097 KVPEFVEPNPLPIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMEGPPGAEPPQ 1153

RESULT 2  
AAB07360  
ID AAB07360 standard; protein; 1153 AA.  
XX AAB07360;  
AC AAB07360;  
XX  
DT 17-JAN-2001 (first entry)  
XX  
DE Human CD11b protein sequence.  
XX  
DE Human; macrophage infiltration inhibition; alpha\_d integrin;  
KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;  
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;  
KW atherosclerosis; multiple sclerosis; asthma; psoriasis; beta2 integrin;  
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;  
KW rheumatoid arthritis; central nervous system injury; CD11b.

OS Homo sapiens.  
XX  
XX WO200029446-A1.  
XX 25-MAY-2000.  
XX  
XX 16-NOV-1999; 99WO-US027139.  
XX  
XX 16-NOV-1998; 98US-00193043.  
XX 08-JUL-1999; 99US-00350259.  
XX  
XX (ICOS-) ICOS CORP.  
XX  
XX Gallatin MW, Van Der Vieren M;  
XX  
XX WPI; 2000-387751/33.

Use of novel anti-alpha integrin d monoclonal antibodies to inhibit  
macrophage infiltration and reduce inflammation at central nervous system  
injury sites.

Example 5; Fig 1; 270pp; English.

Integrins are a class of membrane-associated molecules that participate  
in cellular adhesion. Integrins are made up of an alpha subunit and a  
beta subunit. One class of human integrins are restricted to expression  
in white blood cells and have a common beta2 subunit: the leukocyte  
integrins, Leu-CAMs, leukointegrins or beta2 integrins. Beta2 integrins  
have an important role in immune and inflammatory responses. The present  
protein sequence is the human integrin alpha subunit CD11b. This sequence  
was used in an alignment to identify a novel beta2 integrin alpha  
subunit: alpha\_d (ZAA60014 and AAB07359). The present sequence has  
approximately 60% identity to the protein sequence of alpha\_d. The  
Alpha\_d gene and protein may be useful in therapy for diseases linked to  
aberrant alpha\_d function e.g. Type I diabetes, atherosclerosis, multiple  
sclerosis, asthma, psoriasis, lung inflammation, acute respiratory  
distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency  
(LAD). In addition, anti-alpha\_d monoclonal antibodies may be used in the  
inhibition of macrophage infiltration at the site of a central nervous  
system injury. The monoclonal antibodies can also be used to detect and  
diagnose Crohn's disease

1021	FFGIQEFNATLKNLSFDWYIKTSHNHLINVTABILFNDVSFTLLPGQCAFVRSQTET	1080
1037	FFGIQEFNATLKNLSFDWYIKTSHNHLINVTABILFNDVSFTLLPGQCAFVRSQTET	1096
1081	KVEPFVFPNPLPLIVGSSVGLLLALITAAALYKLGFFKROYKDMSEGPPGAEPQ	1137
1097	KVEPFVFPNPLPLIVGSSVGLLLALITAAALYKLGFFKROYKDMSEGPPGAEPQ	1153
RESULT 3		
AAU80252		
ID	AAU80252 standard; protein; 1153 AA.	
XX	AAU80252;	
AC	AAU80252;	
DT	15-JUL-2002 (first entry)	
XX	Human integrin 1 alpha-M subunit protein.	
DE	Integrin; antiinflammatory; immunosuppression; nephritis; dermatitis;	
XX	Inflammatory disease; autoimmune disease; Crohn's disease;	
KW	human immunodeficiency virus; HIV; myocardial infarction;	
KW	Sjorgen's syndrome; rheumatoid arthritis.	
XX	Homo sapiens.	
OS		
XX	Key	Location/Qualifiers
FH	Misc-difference	/note="Encoded by GGG CAG AGG"
FT	499.500	
XX	WC200218583-A2.	
PN	07-MAR-2002.	
XX	31-AUG-2001; 2001WO-US027227.	
PD		
XX	01-SEP-2000; 2000US-0229700P.	
XX	(BLOO-) CENT BLOOD RES INC.	
PR		
XX	Springer TA, Shimooka M, Lu C;	
PA	WPI; 2002-382964/41.	
PI	N-PSDB; ABK50046.	
XX	Modified integrin-I or integrin I-like domain polypeptide useful as an	
XX	immunogen to produce antibodies specific to polypeptide, comprises a	
PT	disulfide bond such that polypeptide is stabilized in a desired	
PT	conformation.	
XX	Disclosure; Page 109-112; 112pp; English.	
PS		
XX	This invention relates to a modified integrin-I or integrin I-like domain	
CC	polypeptide comprising at least one disulfide bond so that the domain is	
CC	stabilised in a desired conformation. The polypeptide of the invention	
CC	may have antiinflammatory or immunosuppressive activities. The	
CC	polypeptides of the invention have an open conformation and are useful as	
CC	immunogens to produce antibodies that selectively bind to integrin I-	
CC	domain; and for identifying a modulator of integrin activity, or of	
CC	interaction of an integrin and a cognate ligand. The polypeptide of the	
CC	invention, or antibodies (preferably anti-LFA-1 antibody) is useful for	
CC	treating or preventing an integrin mediated disorder which is an	
CC	inflammatory or autoimmune disorder in a subject and for inhibiting the	
CC	binding of an integrin to a cognate ligand such as Crohn's disease,	
CC	nephritis; human immunodeficiency virus (HIV), myocardial infarction,	
CC	Sjorgen's syndrome, rheumatoid arthritis, dermatitis. A therapeutic	
CC	composition comprising the peptide of the invention is useful for	
CC	treating an integrin mediated disorder in a subject. The polypeptides	
CC	and/or active or antigenic fragments are useful as reagents for diagnosis	
CC	of integrin-mediated disorders. The present sequence represents the human	
CC	integrin-1 alpha-M protein subunit used to generate the mutant	
CC	polypeptides of the invention	

[illegible]

SQ		Sequence 1153 AA;	
Query Match		99.9%; Score 5868; DB 5; Length 1153;	
Best Local Similarity		99.8%; Pred. No. 0;	
Matches 1133; Conservative		3; Mismatches 1; Indels 0; Gaps 0;	
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGSLYQCDYSTGSCPEI	60
DB	17	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAANORGSLYQCDYSTGSCPEI	76
QY	61	RLOQPVAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK	120
DB	77	RLOQPVAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQPOK	136
QY	121	FPEALRGCCQEDSDIAFLDGSIIIPHDFRRMKFVSTVMEQLKSKTLFSLMOYSEEF	180
DB	137	FPEALRGCCQEDSDIAFLDGSIIIPHDFRRMKFVSTVMEQLKSKTLFSLMOYSEEF	196
QY	181	RIHFTFEKQNNPNRSLVKPIITQLLGRTHATGVRKVIKRELNITNGARKNAFKILIVI	240
DB	197	RIHFTFEKQNNPNRSLVKPIITQLLGRTHATGIRKVVRELNIINGARKNAFKILIVI	256
QY	241	TDEKEFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIAKPPRDHVFQVN	300
DB	257	TDEKEFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIAKPPRDHVFQVN	316
QY	301	NFEALKTIONLREKIFAIEGTQTSSSSFEHMSQEGFSAITNSGPIILSTVGSYDWAG	360
DB	317	NFEALKTIONLREKIFAIEGTQTSSSSFEHMSQEGFSAITNSGPIILSTVGSYDWAG	376
QY	361	GVFLYTSKEKSFINNTRVDSNDMDAYLGAAAILLRNRVQSLVGLGAPYOHIGLVAMPR	420
DB	377	GVFLYTSKEKSFINNTRVDSNDMDAYLGAAAILLRNRVQSLVGLGAPYOHIGLVAMPR	436
QY	421	QNTGMWESANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSVCL	480
DB	437	QNTGMWESANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTREGQVSVCL	496
QY	481	PRGORARWOCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	540
DB	497	PRGORARWOCDAVLYGEOGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF	556
QY	541	HGTSGSIGSPSHORSIAGSKLSPRLQVFGOSGGQDLTMDGLVDLTGAGQHVILLRSQ	600
DB	557	HGTSGSIGSPSHORSIAGSKLSPRLQVFGOSGGQDLTMDGLVDLTGAGQHVILLRSQ	616
QY	601	PVLVRKAIMFNPREVARNVFECDQVWKGAGEVRVCLHVQKSTRDLREGQIQSVVT	660
DB	617	PVLVRKAIMFNPREVARNVFECDQVWKGAGEVRVCLHVQKSTRDLREGQIQSVVT	676
QY	661	YDLALDSGRPHSAVENETKNSRRTOVLGLTQTCETLKLQIPNCIEDPVSIVLRINF	720
DB	677	YDLALDSGRPHSAVENETKNSRRTOVLGLTQTCETLKLQIPNCIEDPVSIVLRINF	736
QY	721	SLVGTPLSAFGLNLRPVLAEDAQRLFTALPPFEKNCNDNICODDLGITFSFMSLDCLVVG	780
DB	737	SLVGTPLSAFGLNLRPVLAEDAQRLFTALPPFEKNCNDNICODDLGITFSFMSLDCLVVG	796
QY	781	GPREFNVTVVRNDGEDSVRTQVTFPPDLDSVRKYSTIQNQRSQSRWLACASSTEV	840
DB	797	GPREFNVTVVRNDGEDSVRTQVTFPPDLDSVRKYSTIQNQRSQSRWLACASSTEV	856
QY	841	SGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANTYSENMPRTNKTEF	900
DB	857	SGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANTYSENMPRTNKTEF	916
QY	901	QLELPVKYAVYVMTVSHGVSSTKYNFTASENTSRVMQHOVQVSNLQORSLSPISLVFLPV	960
DB	917	QLELPVKYAVYVMTVSHGVSSTKYNFTASENTSRVMQHOVQVSNLQORSLSPISLVFLPV	976
QY	961	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP	1020
DB	977	RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP	1036
QY	1021	FFGIQEEFNATLKGNLSDFWYIKTSHNHLIIYSTABILFNDSVFTLLPGOGAFVRSQTET	1080
DB	1037	FFGIQEEFNATLKGNLSDFWYIKTSHNHLIIYSTABILFNDSVFTLLPGOGAFVRSQTET	1096
QY	1081	KVEPFEVPNPPLPIVGVSSVGGILLALITAAALYKLGFFRQYKDMMSGGPPGAEQ	1137
DB	1097	KVEPFEVPNPPLPIVGVSSVGGILLALITAAALYKLGFFRQYKDMMSGGPPGAEQ	1153
RESULT 5			
AAO14428			
ID	AAO14428 standard; protein; 1153 AA.		
XX	AAO14428;		
AC	AAO14428;		
XX	03-MAY-2002 (first entry)		
DT	Integrin Mac-1 alpha subunit.		
XX	Mac-1; integrin alpha subunit; variant integrin inserted domain protein;		
DE	open conformation; integrin related inflammatory disorder;		
XX	integrin related immunological disorder; rheumatoid arthritis; ischaemia;		
KW	reperfusion; hypovolemic shock; infarction; cerebral shock;		
KW	viral infection; cancer; gene therapy; vaccine;		
KW	bioactive agent screening.		
XX	Unidentified.		
OS	WO200204521-A2.		
XX	17-JAN-2002.		
PN	09-JUL-2001; 2001WO-US021805.		
XX	07-JUL-2000; 2000US-0216600P.		
XX	(CALY ) CALIFORNIA INST OF TECHNOLOGY.		
PA	(BLOO-) CENT BLOOD RES.		
XX	Springer T;		
PI	WPI; 2002-148167/19.		
XX	New integrin I domain protein having alteration in at least 2		
DR	noncontiguous regions and exits in an open conformation, useful for		
XX	treating, preventing or suppressing inflammatory or immunological		
PT	disorders.		
PT	Example 1; Fig 1F; 90pp; English.		
XX	The invention comprises structurally biased variant integrin inserted (1)		
CC	domain proteins, wherein the alterations to the protein occur in at least		
CC	two noncontiguous regions. Specifically the variant integrin I domain		
CC	proteins are structurally biased to exist in the open conformation,		
CC	thereby altering the binding ability of the protein. The invention also		
CC	comprises nucleic acids encoding the variant integrin I domain proteins.		
CC	The integrin I domain proteins and nucleic acids are useful for treating,		
CC	preventing or suppressing integrin related inflammatory and immunological		
CC	disorders (e.g. rheumatoid arthritis). The variant integrin I domain		
CC	proteins and nucleic acids can also be used for treating: ischaemia/		
CC	reperfusion (e.g. hypovolemic shock); infarction; cerebral shock; viral		
CC	infection; and cancer. The variant integrin I domain nucleic acids and		
CC	proteins may be used in gene therapy, as vaccines and to screen for		
CC	bioactive agents. The present amino acid sequence represents the Mac-1		
CC	alpha subunit of integrin		
XX	Sequence 1153 AA;		
QY	Query Match 99.9%; Score 5868; DB 5; Length 1153;		
DB	Best Local Similarity 99.6%; Pred. No. 0;		
DB	Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;		

QY 1 FNLDTENAMTFQENARGFGOSVVOLOQSRVVVGAPOBIIVAAORGSLYQCDYSTGSCPEI 60  
DB |||||  
QY 17 FNLDTENAMTFQENARGFGOSVVOLOQSRVVVGAPOBIIVAAORGSLYQCDYSTGSCPEI 76  
DB |||||  
QY 61 RLQVPVEAVNMSLGLSLAATTPSPOLLACGPTVHQTCSNTYVKGCLFGLFGSNLRQOPQK 120  
DB |||||  
QY 77 RLQVPVEAVNMSLGLSLAATTPSPOLLACGPTVHQTCSNTYVKGCLFGLFGSNLRQOPQK 136  
DB |||||  
QY 121 FPEALRCPOEDSDIAFLIDGSGSIIIPHDRRMKEFVSTWMEQLKSKTILFSLMOYSEEF 180  
DB |||||  
QY 137 FPEALRCPOEDSDIAFLIDGSGSIIIPHDRRMKEFVSTWMEQLKSKTILFSLMOYSEEF 196  
DB |||||  
QY 181 RIHPTFKFQNNPNSRLVPIITOLLGRTHATGVKRVIRELNIITNGARKNAFKILIVI 240  
DB |||||  
QY 197 RIHPTFKFQNNPNSRLVPIITOLLGRTHATGVKRVIRELNIITNGARKNAFKILIVI 256  
DB |||||  
QY 241 TDGEKFGDPIGYEDVPEADREGVIRVYGVGDAPRSEKSRQELNTIASKPPRDHVPQVN 300  
DB |||||  
QY 257 TDGEKFGDPIGYEDVPEADREGVIRVYGVGDAPRSEKSRQELNTIASKPPRDHVPQVN 316  
DB |||||  
QY 301 NFEALKTIONQIREKIEFAIGTGTGSSSSPEHEMSQEGFSAITNSGPLLSTVGSYDWAG 360  
DB |||||  
QY 317 NFEALKTIONQIREKIEFAIGTGTGSSSSPEHEMSQEGFSAITNSGPLLSTVGSYDWAG 376  
DB |||||  
QY 361 GVFLYTSKESKSTFNMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420  
DB |||||  
QY 377 GVFLYTSKESKSTFNMTRVDSMDNDAYLGYAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 436  
DB |||||  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSCPL 480  
DB |||||  
QY 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYEQTRGGQVSCPL 496  
DB |||||  
QY 481 PRGORARWQCDVLYGQGGPWGFGAALTJVLGVNGDKLTDVAIGAPGEEDNRGAYILF 540  
DB |||||  
QY 497 PRGORARWQCDVLYGQGGPWGFGAALTJVLGVNGDKLTDVAIGAPGEEDNRGAYILF 556  
DB |||||  
QY 541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 600  
DB |||||  
QY 557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTGVAQGHVLLRSQ 616  
DB |||||  
QY 601 PVLRVKAIMEFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB |||||  
QY 617 PVLRVKAIMEFNPREVARNVPECDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
DB |||||  
QY 661 YDLALDSGRSHRAVFNETKNSRQTVLGTCTETKLQPLNCIEDPVSPTVLRLNF 720  
DB |||||  
QY 677 YDLALDSGRSHRAVFNETKNSRQTVLGTCTETKLQPLNCIEDPVSPTVLRLNF 736  
DB |||||  
QY 721 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITTFMSLDCLVVG 780  
DB |||||  
QY 737 SLVGTPLSAFGLNLRPVLAEADAQRLFTALFPFEKNCNDNICQDDLSITTFMSLDCLVVG 796  
DB |||||  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVYKVKSTLQNRQSQRWRLACESASSTEV 840  
DB |||||  
QY 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLVYKVKSTLQNRQSQRWRLACESASSTEV 856  
DB |||||  
QY 841 SGALKSTCSINHPFIPENSEVFNITPDVDSKASLGNKLLKANVTSENNMPTNKTET 900  
DB |||||  
QY 857 SGALKSTCSINHPFIPENSEVFNITPDVDSKASLGNKLLKANVTSENNMPTNKTET 916  
DB |||||  
QY 901 QLELPVKYAVYVWVTVSHGVSTKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPV 960  
DB |||||  
QY 917 QLELPVKYAVYVWVTVSHGVSTKYLNFTASENTSRVMOHQVSNLQORSPLISLVFLVPV 976  
DB |||||  
QY 961 RLNQTIVDRPQVTFNSNLSTCHTKERLPSSHDFLAEKAPVNVNCSIAVCQRIQCDIP 1020  
DB |||||  
QY 977 RLNQTIVDRPQVTFNSNLSTCHTKERLPSSHDFLAEKAPVNVNCSIAVCQRIQCDIP 1036  
DB |||||  
QY 1021 FFGIQEEFNATLKGNLFDWYIKTSHNHLIVSTAEILFNDVSTLTPGGQAFVRSOTET 1080  
DB |||||  
QY 1037 FFGIQEEFNATLKGNLFDWYIKTSHNHLIVSTAEILFNDVSTLTPGGQAFVRSOTET 1096  
DB |||||  
QY 1081 KVEPFEVFNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKQKDMNSEGGPPGAEPQ 1137

DB 1097 KVEPFEVFNPLPLIVGSSVGLLILALITAAALYKLGFFKQYKQKDMNSEGGPPGAEPQ 1153  
|||  
RESULT 6  
ADD25615  
ID ADD25615 standard; protein; 1153 AA.  
XX AC ADD25615;  
XX DT  
XX 15-JAN-2004 (first entry)  
XX Binding domain-immunoglobulin fusion protein-associated protein #85.  
XX Binding domain; immunoglobulin; fusion protein; cytostatic;  
XX antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
XX neuroprotective; hinge region; immunoglobulin heavy chain;  
XX CH2 constant region; CH3 constant region; IgG1;  
XX antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
XX malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
XX rheumatoid arthritis; myasthenia gravis; Grave's disease;  
XX type I diabetes mellitus; multiple sclerosis; autoimmune disease.  
XX Unidentified.  
OS US2003118592-A1.  
XX PN  
XX 26-JUN-2003.  
PD  
XX 25-JUL-2002; 2002US-00207655.  
PF  
XX 17-JAN-2001; 2001US-0367358P.  
PR  
XX 17-JAN-2002; 2002US-0003530.  
PR  
XX 03-JUN-2002; 2002US-0385691P.  
XX (GENE-) GENE-CRAFT INC.  
PA  
XX Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
PI WPI; 2003-801317/75.  
DR  
XX New binding domain-immunoglobulin fusion protein, useful for treating a  
PT subject having or suspected of having a malignant condition or a B-cell  
PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.  
XX  
PS Disclosure; SEQ ID NO 176; 157pp; English.  
XX  
CC The invention relates to a binding domain-immunoglobulin fusion protein  
CC comprising a binding domain polypeptide that is fused to an  
CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
CC CH2 constant region polypeptide that is fused to the hinge region  
CC polypeptide, and an immunoglobulin heavy chain CH3 constant region  
CC polypeptide that is fused to the CH2 constant region polypeptide. The  
CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
CC region polypeptide, derived from (a) having 3 or more cysteine residues;  
CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
CC immunoglobulin hinge region polypeptide contains no more than one  
CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
CC polypeptide, derived from (a) having 3 or more cysteine residues; where  
CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
CC capable of at least one immunological activity comprising antibody  
CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
CC binding domain polypeptide is capable of specifically binding to an  
CC antigen. Also included are an isolated polynucleotide encoding the  
CC binding domain-immunoglobulin fusion protein, a recombinant expression  
CC construct comprising the polynucleotide (operably linked to a promoter),  
CC a host cell transformed or transfected with a recombinant expression  
CC construct, producing the binding domain-immunoglobulin fusion protein, a



CC pharmaceutical composition comprising the binding domain-immunoglobulin  
 CC fusion protein or polynucleotide and a carrier, and treating a subject  
 CC having or suspected of having a malignant condition or a B-cell disorder.  
 CC The binding domain-immunoglobulin fusion protein is useful for treating a  
 CC subject having or suspected of having a malignant condition or a B-cell  
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
 CC myasthenia gravis, Grave's disease, type I diabetes mellitus, multiple  
 CC sclerosis or autoimmune disease. The present sequence is a binding domain  
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
 CC sequence data for this patent formed part of the printed specification  
 CC and is also available in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030118592. The authors have not  
 CC identified the sequences in the printed specification by their SEQ ID  
 CC number therefore none of the sequences can be explicitly identified.  
 XX  
 SQ Sequence 1153 AA;

Query Match 99.9%; Score 5868; DB 7; Length 1153;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVLQGSRRVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 60  
 DB 17 FNLDTENAMTFQENARGFGQSVVLQGSRRVVGAPQBIVAANQSGSLYQCDYSTGSCPEI 76  
 QY 61 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQ 120  
 DB 77 RLQVPVEAVNMSLGLSLAATTPPOLLACGPTVHOTCSENTYVKGCLFLGSLNRQOPQ 136  
 QY 121 FPALRGCPQEDSDIAPLIDGSGIIPHPFRMKFVSTVMEQLKSKTLFSLMQYSEEF 180  
 DB 137 FPALRGCPQEDSDIAPLIDGSGIIPHPFRMKFVSTVMEQLKSKTLFSLMQYSEEF 196  
 QY 181 RIHFTKEFONNPNRSLVKPIQLLGRTHATATGVRKVIPELLNITNGARKNAFKILVI 240  
 DB 197 RIHFTKEFONNPNRSLVKPIQLLGRTHATATGVRKVIPELLNITNGARKNAFKILVI 256  
 QY 241 TDEKEGDPGLGYEDVPEADREGVIRVIGVGAFAFRSEKSRQELNTIASKPPRDHVFQVN 300  
 DB 257 TDEKEGDPGLGYEDVPEADREGVIRVIGVGAFAFRSEKSRQELNTIASKPPRDHVFQVN 316  
 QY 301 NFPAKTIQNLREKIFAIEGTQTGSSSPSEHEMSQEGFSAAITSNGLLSTVGSYDWAG 360  
 DB 317 NFPAKTIQNLREKIFAIEGTQTGSSSPSEHEMSQEGFSAAITSNGLLSTVGSYDWAG 376  
 QY 361 GVFLYTSKEKSTFINTRVDSMDNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 420  
 DB 377 GVFLYTSKEKSTFINTRVDSMDNDAYLGYAAAIILNRVQSLVGLAPRYQHIGLVAMFR 436  
 QY 421 QNTGMWESANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGGQVSCPL 480  
 DB 437 QNTGMWESANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTGRGGQVSCPL 496  
 QY 481 PRQGRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
 DB 497 PRQGRARWQCDVLYGEQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
 QY 541 HGTSSGSI SPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 600  
 DB 557 HGTSSGSI SPHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAGQHVLLRSQ 616  
 QY 601 PVLVRKAIIMEFNPREVARNVFECDQVVGKEAGEVRVCLHVOKSTRDLREGQIOSVVT 660  
 DB 617 PVLVRKAIIMEFNPREVARNVFECDQVVGKEAGEVRVCLHVOKSTRDLREGQIOSVVT 676  
 QY 661 YDLALDSGRPHSAVFNETKNSTRQTQVLGLTQTCETLKLQPLNCIEDPVSIVLRLNF 720  
 DB 677 YDLALDSGRPHSAVFNETKNSTRQTQVLGLTQTCETLKLQPLNCIEDPVSIVLRLNF 736  
 QY 721 SLVGTPLSAFNGRLPVLAEADAQLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCILVWG 780  
 DB 737 SLVGTPLSAFNGRLPVLAEADAQLFTALFPFPEKNCNDNICQDDLSITFSFMSLDCILVWG 796

QY 781 GPREENVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACSSASSTEV 840  
 DB 797 GPREENVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNSORSWRLACSSASSTEV 856  
 QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900  
 DB 857 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 916  
 QY 901 QLELPVKIAYVMVTVSHGVSTKYLVNTASENTSRVMOHQYQVSNLQORSPLISLVFLVPV 960  
 DB 917 QLELPVKIAYVMVTVSHGVSTKYLVNTASENTSRVMOHQYQVSNLQORSPLISLVFLVPV 976  
 QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQDIP 1020  
 DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQDIP 1036  
 QY 1021 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGCGAFVRSQTE 1080  
 DB 1037 FFGIQEEFNATLKNLSFDWYIKTSHNHLIIVSTAELFNDSVFTLLPGCGAFVRSQTE 1096  
 QY 1081 KVTFPEVPNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1137  
 DB 1097 KVTFPEVPNPPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPQ 1153  
 RESULT 7  
 ID AAR04136 standard; protein; 1153 AA.  
 XX AAR04136;  
 XX AC AC  
 DT 09-SEP-2004 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 07-SEP-1990 (first entry)  
 XX  
 DE Alpha subunit of Mac-1 leukocyte adhesion receptor.  
 XX  
 KW Mac-1 alpha subunit; Mac-1 alpha/beta heterodimer;  
 KW non-specific defence system; integrin gene superfamily.  
 XX  
 OS Synthetic.  
 FH Key  
 FT Region 1..116  
 FT /label= signal\_peptide  
 FT Modified-site 86..88  
 FT /label= putative N-glycosylation site  
 FT Modified-site 240..242  
 FT /label= putative N-glycosylation site  
 FT Modified-site 391..393  
 FT /label= putative N-glycosylation site  
 FT Modified-site 469..471  
 FT /label= putative N-glycosylation site  
 FT Modified-site 693..695  
 FT /label= putative N-glycosylation site  
 FT Modified-site 697..699  
 FT /label= putative N-glycosylation site  
 FT Modified-site 735..737  
 FT /label= putative N-glycosylation site  
 FT Modified-site 802..804  
 FT /label= putative N-glycosylation site  
 FT Modified-site 881..883  
 FT /label= putative N-glycosylation site  
 FT Modified-site 901..903  
 FT /label= putative N-glycosylation site  
 FT Modified-site 912..914  
 FT /label= putative N-glycosylation site  
 FT Modified-site 941..943  
 FT /label= putative N-glycosylation site  
 FT Modified-site 947..949  
 FT /label= putative N-glycosylation site  
 FT Modified-site 979..981  
 FT /label= putative N-glycosylation site

FT Modified-site 994..996  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1022..1024  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1045..1047  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1051..1053  
 FT /label= putative N-glycosylation site  
 FT Modified-site 1076..1078  
 FT /label= putative N-glycosylation site  
 FT Region 1106..1134  
 FT /label= putative\_transmembrane\_region  
 XX  
 PN EP364690-A.  
 XX  
 XX 25-APR-1990.  
 XX  
 XX 17-AUG-1989; 89EP-00115159.  
 XX  
 XX 23-AUG-1988; 88US-002315353.  
 PR 09-MAR-1989; 89US-00321239.  
 XX  
 XX (DAND ) DANA FARBER CANCER INST INC.  
 FA  
 XX Springer TA, Corbi A;  
 PI  
 XX WPI; 1990-125938/17.  
 DR N-PDSB; AAQ04043.  
 DR  
 XX  
 FT New pure Mac-1 alpha sub-unit and corresp. DNA - useful for treating  
 FT inflammation and viral infections, and in diagnosis.  
 FT  
 XX Disclosure; Page ?; 3pp; English.  
 XX  
 CC Mac-1 alpha subunit is involved in the response to inflammation, i.e.  
 CC recognition of and migration to sites of inflammation. It also attaches  
 CC to cellular substrates as part of this function making it useful in  
 CC visualising endothelial tissue. Mac-1 is a member of the Integrin Gene  
 CC superfamily. (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25  
 CC -MAR-2003 to correct PA field.)  
 CC  
 CC Revised record issued on 09-SEP-2004 : Correction to feature table key  
 CC and pages  
 CC  
 XX Sequence 1153 AA;  
 SQ  
 Query Match 99.7%; Score 5858; DB 2; Length 1153;  
 Best Local Similarity 99.6%; Pred. No. 0;  
 Matches 1132; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 ENLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI 60  
 DB 17 ENLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI 76  
 QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHTQCSNTYVKGCLFLFGLNLRQQPQK 120  
 DB 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHTQCSNTYVKGCLFLFGLNLRQQPQK 136  
 QY 121 PPEALRGCPQSDSDIAFLIDGSGIIPHDFFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 180  
 DB 137 PPEALRGCPQSDSDIAFLIDGSGIIPHDFFRMKEFVSTVMEQLKSKTLFSLMQYSEEF 196  
 QY 181 RIHFTTFEQQNPNSRSLVKPITQLLGRTHATGVRKVIRELNTNGARKNAFKILIVI 240  
 DB 197 RIHFTTFEQQNPNSRSLVKPITQLLGRTHATGVRKVIRELNTNGARKNAFKILIVI 256  
 QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIAKPPRHHVQVN 300  
 DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIAKPPRHHVQVN 316  
 QY 301 NFEALKTTQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAATNSGPIILSTVGSYDWAG 360  
 DB 317 NFEALKTTQNLREKIFAIEGTQTGSSSPFHEMSQEGFSAATNSGPIILSTVGSYDWAG 376

QY 361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420  
 DB 377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGYAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436  
 QY 421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYQYEQTRGGQVSCPL 480  
 DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYQYEQTRGGQVSCPL 496  
 QY 481 PRQARWQCDVLYGEGQGWGRFGAALTVLGVNKGKLTVDVAIGAGBEDNRGAYLIF 540  
 DB 497 PRQARWQCDVLYGEGQGWGRFGAALTVLGVNKGKLTVDVAIGAGBEDNRGAYLIF 556  
 QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTWVAGHVVLLRSQ 600  
 DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTWVAGHVVLLRSQ 616  
 QY 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRLHVQKSTRDLREGQIQSVVT 660  
 DB 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRLHVQKSTRDLREGQIQSVVT 676  
 QY 661 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 720  
 DB 677 YDLALDSGRPHSRAVFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPSPVILRLNF 736  
 QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 780  
 DB 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFKKNCNDNICODDLSITFSFMSLDCLVVG 796  
 QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 840  
 DB 797 GPREFNVTVVRNDGEDSYRTQVTFPPPLDLSYRKVSTLQNRQSRWRLACESASSTEV 856  
 QY 841 SGALKSTSCSINHPIFENSEVTEITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
 DB 857 SGALKSTSCSINHPIFENSEVTEITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916  
 QY 901 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHYQVSNLQSRSLPISLVFLVPV 960  
 DB 917 QLELPVKYAVVMVTVSHGVSTKYLNFTASENTSRVMQHYQVSNLQSRSLPISLVFLVPV 976  
 QY 961 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCRIQCIP 1020  
 DB 977 RLNQTIVDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCRIQCIP 1036  
 QY 1021 PFGIQEENATLKGNSLSPDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQGAFVRSQTEF 1080  
 DB 1037 PFGIQEENATLKGNSLSPDWYIKTSHNHLIIVSTAEILFNDVSFTLLPQGAFVRSQTEF 1096  
 QY 1081 KVEPFEPVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137  
 DB 1097 KVEPFEPVNPPLIVGSSVGGILLALITAAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153  
 RESULT 8  
 ADM9589  
 ID ADM9589 standard; protein; 1152 AA.  
 XX  
 AC ADM9589;  
 XX  
 DT 17-JUN-2004 (first entry)  
 XX  
 DE Human integrin alphaM subunit precursor protein.  
 XX  
 KW integrin alpha subunit; beta; antiposrotic; thrombolytic; anticoagulant;  
 KW osteopathic; cytosolic; immunosuppressive; anti-inflammatory;  
 KW neuroprotective; antisticking; immunotherapy; inflammatory;  
 KW autoimmune disorder; thrombosis; cancer; osteoporosis;  
 KW sickle cell anaemia; psoriasis; multiple sclerosis; human; precursor;  
 KW alpham.  
 XX  
 OS Homo sapiens.  
 XX

Key	Location/Qualifiers	
FT	Misc-difference 965	
FT	/note= "Encoded by CCC"	
XX		
PN	WO2004007530-A2.	
PD	22-JAN-2004.	
XX		
XX	17-JUL-2003; 2003WO-US022301.	
XX	17-JUL-2002; 2002US-0396783P.	
PR	17-JUL-2002; 2002US-0396790P.	
PR	11-SEP-2002; 2002US-0410135P.	
XX		
PA	(BLOO-) CENT BLOOD RES INC.	
XX		
PI	Springer TA, Takagi J;	
XX	WPI; 2004-122877/12.	
DR	N-PSDB; ADM9598.	
XX		
PT	Novel modified integrin protein having extracellular domains of integrin alpha and beta subunits or integrin alpha1 and beta3 subunit, useful for treating integrin mediated disorders.	
PT		
PT		
XX		
PS	Disclosure; SEQ ID NO 4; 232pp; English.	
XX		
CC	The invention relates to a novel isolated or recombinant modified integrin protein having extracellular domains of integrin alpha and beta subunits where one of the subunits has one or more mutations, an altered surface feature or an amino acid substitution or internal deletion, extracellular domains of the integrin beta subunit that comprise a mutation that alters a non-cysteine residue to cysteine or extracellular domains of integrin alpha and beta subunits. The polypeptide of the invention demonstrates antipruritic, thrombolytic, anticoagulant, osteoprotective, cytotatic, immunosuppressive, antiinflammatory, neuroprotective and anticikling activities and may be useful for immunotherapy in order to prevent or treat an integrin-mediated disorder such as an inflammatory disorder, an autoimmune disorder, thrombosis, cancer, osteoporosis, sickle cell anaemia, psoriasis and multiple sclerosis. The current sequence is that of the human integrin alpha subunit precursor protein of the invention.	
CC		
XX		
SQ	Sequence 1152 AA;	
Query Match 99.6%; Score 5852.5; DB 8; Length 1152;		
Best Local Similarity 99.6%; Pred. No. 0;		
Matches 1132; Conservative 3; Mismatches 1; Indels 1; Gaps 1;		
QY	1 FNLDTENAMTFQBNARGFGQSVVQLOGSRVVVGAPQEIIVAANORGSLYQCDYSTGSCEPI 60	
DB	17 FNLDTENAMTFQBNARGFGQSVVQLOGSRVVVGAPQEIIVAANORGSLYQCDYSTGSCEPI 76	
QY	61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGNSLNRQOPQK 120	
DB	77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGCLFLGNSLNRQOPQK 136	
QY	121 FPEALGCGQESDIAFLIDGSGSIIPHDFRMRKEFVSTVMEOLKSKTFLSLMQYSEEF 180	
DB	137 FPEALGCGQESDIAFLIDGSGSIIPHDFRMRKEFVSTVMEOLKSKTFLSLMQYSEEF 196	
QY	181 RIHFTKEFQNNPNRSLVKPIITQLLGRTHATGVRKPIRELLNITNGARKNAFKILVI 240	
DB	197 RIHFTKEFQNNPNRSLVKPIITQLLGRTHATGIRKVRLEFNITNGARKNAFKILVI 256	
QY	241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQYN 300	
DB	257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPPRDHVFQYN 316	
QY	301 NFPEALTIQNLREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG 360	
DB	317 NFPEALTIQNLREKIFAIEGTQTGSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG 376	
QY	361 GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYOHTGLVAMER 420	
DB	377 GVFLYTSKSKSTFINMTRVDSMDNDAYLGAAAIILNRNVSQSLVGLAPRYOHTGLVAMER 436	
QY	421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSCPL 480	
DB	437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYYEOTRGQVSCPL 496	
QY	481 PRGQARWOCDAVLYGEOGQPGWRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYLIF 540	
DB	497 PRG-RARWOCDAVLYGEOGQPGWRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYLIF 555	
QY	541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSSGGDLTMDGLVDLTGVAQGHVLLRSQ 600	
DB	556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGOSLSSGGDLTMDGLVDLTGVAQGHVLLRSQ 615	
QY	601 PVLVRKAIMFEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIOSVVT 660	
DB	616 PVLVRKAIMFEFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRREGQIOSVVT 675	
QY	661 YDLALDSGRPHSRAPVNETKNSTRQTVLGLTQCTETLKLQLPNCIEDPVPVILRLNF 720	
DB	676 YDLALDSGRPHSRAPVNETKNSTRQTVLGLTQCTETLKLQLPNCIEDPVPVILRLNF 735	
QY	721 SLVGTPLSAFAGNLRPVLAEDAQRLEFALPFPEKNCNGNDNICODDLSITTSFMSLDCLVG 780	
DB	736 SLVGTPLSAFAGNLRPVLAEDAQRLEFALPFPEKNCNGNDNICODDLSITTSFMSLDCLVG 795	
QY	781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSLACESASSTEV 840	
DB	796 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSYRKVSTLQNRQSRWSLACESASSTEV 855	
QY	841 SGALKSTCSINHIPIPPENSEVTNITFDVDSKASLGKLLKLLKANTSENMPRTNKTEF 900	
DB	856 SGALKSTCSINHIPIPPENSEVTNITFDVDSKASLGKLLKLLKANTSENMPRTNKTEF 915	
QY	901 QLELPVKYAVYVMVTSHGVSSTKYNFTASNTSRVMQHOYQVSNLQORSIPISLAVFLVPV 960	
DB	916 QLELPVKYAVYVMVTSHGVSSTKYNFTASNTSRVMQHOYQVSNLQORSIPISLAVFLVPV 975	
QY	961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1020	
DB	976 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1035	
QY	1021 FFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTABILFNDSVFTLLPGQGFVRSQTET 1080	
DB	1036 FFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTABILFNDSVFTLLPGQGFVRSQTET 1095	
QY	1081 KVEPPEVPNPLPLIYGVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEFQ 1137	
DB	1096 KVEPPEVPNPLPLIYGVSSVGGLLLLALITAAALYKLGFFKQYKDMSEGPPGAEFQ 1152	
RESULT 9		
ADP12435		
ID	ADP12435 standard; protein; 1152 AA.	
XX	ADP12435;	
AC	ADP12435;	
XX		
DT	12-AUG-2004 (first entry)	
XX		
DE	Protein encoded by mRNA of the invention #45.	
XX		
KW	transplant rejection; immune system; rheumatoid arthritis; lupus;	
KW	inflammatory bowel disease; multiple sclerosis; HIV; AIDS.	
OS	Homo sapiens.	
XX		
PN	WO2004042346-A2.	
XX		
PD	21-MAY-2004.	
XX		
PF	24-APR-2003; 2003WO-US012946.	

XX 24-APR-2002; 2002US-00131831.  
PR 20-DEC-2002; 2002US-00325899.  
XX (EXPR-) EXPRESSION DIAGNOSTICS INC.  
PA Wohlgemuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;  
PI Rosenberg S;  
PI  
XX WPI; 2004-400724/37.  
DR  
XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
PT rejection, in an individual, comprises detecting the expression level of  
PT the genes.  
XX  
PS Claim 65; SEQ ID NO 2444; 1762pp; English.  
XX  
CC The present invention relates to diagnosing or monitoring transplant  
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual  
CC comprises detecting the expression level of one or more genes. The  
CC methods, system and kits are useful in diagnosing or monitoring  
CC islet, lung, bone marrow or stem cell transplant rejection,  
CC xenotransplant rejection or mechanical organ replacement rejection, in an  
CC individual. The method is also useful in assessing the immune status of  
CC an individual. The methods are also useful in diagnosing and monitoring  
CC diseases that involve the immune system, e.g. rheumatoid arthritis,  
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
CC viral, bacterial or fungal infection. The present sequence represents a  
CC protein that is encoded by the mRNA of the invention.  
XX  
SQ Sequence 1152 AA;

Query Match 99.5%; Score 5845.5; DB 8; Length 1152;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 1131; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQBIIVANORGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTFOENARGFGQSVVOLQGSRVVVGAPQBIIVANORGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPEAVNMVSLGLSAAATSPQLLACGTVHQTCTENTYVKGCLFLFGSNLRQOPQK 120  
DB 77 RLQVPEAVNMVSLGLSAAATSPQLLACGTVHQTCTENTYVKGCLFLFGSNLRQOPQK 136  
QY 121 PPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTKTLFSLMQYSEEF 180  
DB 137 PPEALRGCPQSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTKTLFSLMQYSEEF 196  
QY 181 RIHETFEFQNNPNSRIVKPEITOLLGRTHATGVRKVIKELNITNGARKNAFKILIVI 240  
DB 197 RIHETFEFQNNPNSRIVKPEITOLLGRTHATGVRKVIKELNITNGARKNAFKILIVI 256  
QY 241 TDGEKFGDPLGYEDVIEADREGVIRYVIGVGDAPFRSEKSRQELNITASKPPRPHVQVN 300  
DB 257 TDGEKFGDPLGYEDVIEADREGVIRYVIGVGDAPFRSEKSRQELNITASKPPRPHVQVN 316  
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPFHEMSQGFSAATISNGPLLSVTGVSYDWAG 360  
DB 317 NFEALKTIQNLREKIFAIEGTQTGSSSPFHEMSQGFSAATISNGPLLSVTGVSYDWAG 376  
QY 361 GVFLYTSKESKSTFNMTRVDSMDNDAYLGAAAILLRNQSIVLGAAPRYOHIGLVAMER 420  
DB 377 GVFLYTSKESKSTFNMTRVDSMDNDAYLGAAAILLRNQSIVLGAAPRYOHIGLVAMER 436  
QY 421 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEYQTRGGQVSVCP 480  
DB 437 QNTGWMESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEYQTRGGQVSVCP 496  
QY 481 PRGORARQCDVAVLYGEGQGWGRFGAALTIVLGVNDGDKLTDVAIGAPGEDNRGAVLYF 540  
DB 497 PRG-RARWQCDVAVLYGEGQGWGRFGAALTIVLGVNDGDKLTDVAIGAPGEDNRGAVLYF 555

QY 541 HGTSGSGISPSHSORITAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGQHVLLLRQ 600  
DB 556 HGISGSGISPSHSORITAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTGAGQHVLLLRQ 615  
QY 601 PVLRVKAIMFENPREVARNVFECNDQVVGKGEAGEVRVCLHVKQKSTRDLREGQIQSVVT 660  
DB 616 PVLRVKAIMFENPREVARNVFECNDQVVGKGEAGEVRVCLHVKQKSTRDLREGQIQSVVT 675  
QY 661 YDLALDSGRPHSRAVFNETKNSTRRTQVGLTQTCETLKLQLPNCIEDPVSPIVLINLF 720  
DB 676 YDLALDSGRPHSRAVFNETKNSTRRTQVGLTQTCETLKLQLPNCIEDPVSPIVLINLF 735  
QY 721 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPFKKNCNDNICODDLSITTFGMSLDCLVVG 780  
DB 736 SLVGTPLSAFGNLRPVLAEADAQRLFTALFPFKKNCNDNICODDLSITTFGMSLDCLVVG 795  
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 840  
DB 796 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQSRQSRWLACESASSTEV 855  
QY 841 SGALKSTSCSINHPPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900  
DB 856 SGALKSTSCSINHPPIPPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 915  
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLISLVFLVPV 960  
DB 916 QLELPVKYAVYVMTSHGVSTKYLNFNTASENTSRVMOHQYQVSNLQORSPLISLVFLVPV 975  
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1020  
DB 976 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCQRIQCDIP 1035  
QY 1021 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVETLLPGQAFVRSQDET 1080  
DB 1036 FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDSVETLLPGQAFVRSQDET 1095  
QY 1081 KVPFPFVNPPLIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137  
DB 1096 KVPFPFVNPPLIVGVSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1152  
RESULT 10  
ADP44061  
ID ADP44061 standard; protein; 1163 AA.  
XX  
AC ADP44061;  
XX  
DT 09-SEP-2004 (first entry)  
XX  
DE Human CD11C protein SEQ ID NO:14.  
XX  
KW HIV entry inhibitor; cell surface protein inhibitor; HIV infection;  
KW anti-HIV; virucide; HIV uptake inhibitor; human; CD11C; integrin alpha X.  
XX Homo sapiens.  
XX  
XX WC2004053094-A2.  
XX  
PD 24-JUN-2004.  
XX  
PF 08-DEC-2003; 2003WO-US039208.  
XX  
PR 06-DEC-2002; 2002US-0431522P.  
XX  
XX (PEPP-) PPD DEV LP.  
XX  
XX Dunn SJ, Holzmayer TA;  
XX  
XX WPI; 2004-480928/45.  
XX  
XX N-FSDB; ADP44060.  
XX  
PT Identifying an inhibitor of human immunodeficiency virus (HIV) entry into



```
PT agent for IBD.
XX
PS Example; SEQ ID NO 27; 151pp; Japanese.
XX
CC The invention comprises marker genes for inflammatory bowel disease (IBD)
CC - FcgammaR IIa, FcgammaR IIb, Mig, NKG-2, hexokinase, HM74, REG III,
CC LPAP, Mip-1(beta), L-selectin, EGFL6, IDO, IL-8, CD11c, and TLR2 genes.
CC The DNA and protein sequences of the invention are useful in the
CC detection and treatment of IBD (e.g. ulcerative colitis and Crohn's
CC disease). The present amino acid sequence represents an IBD marker
CC protein of the invention.
XX
SQ Sequence 1163 AA;

Query Match 59.1%; Score 3473; DB 8; Length 1163;
Best Local Similarity 61.3%; Pred. No. 6.9e-280;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFQENARGFGQSVWOLQGSRRVVVGAPQETVAANQRGSLYQCDYSTGSCPEI 60
DB 20 FNLDETELTA PRVDSAGFGDSVVQVANSVVVVGAPQKITAANQTGGLYQCGYSTGACEPI 79

QY 61 RLQVPVAVNMNLSGLSLAATTSPPQLLACGPTVHQTCSNTYVKGLCLFLGNSLRQOPQK 120
DB 80 GLQVPPRAVNMNLSGLSLAATTSPPQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137

QY 121 PPEALRGCPQSDSDIAFLIDGSGTII PHDFRMKEFVSTVMEQLKSKTFLSLMOYSEEP 180
DB 138 LPVSRQECPRQEQDIVFLIDGSGTSSRNFAFMNFVRAVTSOFQRPSTQSLMQFSNKF 197

QY 181 RIHFTFEFONNPRSLVVKPTQLLGRTHATGVKRVIRELLNITNGARKNAKILIVI 240
DB 198 QHTFTFEFRNTSPLSLASVHQLQGFYITATQNVVHRLPHASYGARRDAIKILIVI 257

QY 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGDAPFRSEKSRQELMTIASKPPRDHVFQVN 300
DB 258 TDGRKEGSLDYKDVIPWADAAGIIRYAIYGVLAFQNRNSWKNELNDIAKSPQEHIFKVE 317

QY 301 NFEALKITQNLQREKIFAIEGTQTGSSSSFEHENGQEGFSAITNSGPLLSTVGSYDWAQ 360
DB 318 DFDALKIQNLQREKIFAIEGTETSSSFLEMAQEGFSAVFTPDGVLGAVGFTWSG 377

QY 361 GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAIIILNRVQSLVLCAPRYQHIGLVAMFR 420
DB 378 GAFLYPPNMSPTFINNSGVNDMRDLYGYSTELALWKGVQSLVLCAPRYQHTGKAVFT 437

QY 421 QNTGMWESNANVKGITQICAYFGASLCSDVDVDSNGSTDLVLCAPRYQHTGKAVFT 480
DB 438 QVSRQWRMKAETVGTQIGSYFGASLCSDVDVDSNGSTDLVLCAPRYQHTGKAVFT 497

QY 481 PRQQRARQCDAVLYGEOGPWGRFGAALTVLGVNGDKLTDAVAGPCEEDNRCVAVLF 540
DB 498 PRGWR-RWMCDAVLYGEOGPWGRFGAALTVLGVNGDKLTDAVAGPCEEDNRCVAVLF 556

QY 541 HGTSGSISPSHSORIASGLSPRIQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ 600
DB 557 HGVLPSPISPSHSORIASGLSPRIQYFGQALSGGQDLTQDGLVDLAVGARGQVLLRFR 616

QY 601 PVLRYKALMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVOKSTRDLRLRGQTSVVT 660
DB 617 PVLWVGVSQFIPAEIPRSAPFECREQVSEQLTVQSNICLYIDKRSKLLGSRDLQSSVT 676

QY 661 YDLALDSGPHSRVAFNETKSTRQTVLGTQTCETLKLQLPNCIEPVPSPVILRLNF 720
DB 677 LDALDPGLSPRATFQETKNSLSRVRVILGKAHCENFLLPSCVDSVPTITLRLNF 736

QY 721 SLVGTPLSAFGLNRPVLAEDAQRLEFTALFPFKNCGNDNICODLSITFSFMSLDCLVVG 780
DB 737 TLVGKELLAFNLRPLMALAQRFTASLPFKNCGADHICQNDLIGSFSPGLKSLVG 796

QY 781 GPREFNVTVVRNDGDSRTQVTFEFPDLDLSYKRVSTLQNSQRSWRLACASSTEV 840
DB 797 SNLELNAEVWVNDGDSYGTITTFSHRPAGLSYRYAEGQKQGLRSLHLTCDSPAVG-- 854
```

---

```
QY 841 SGALKSTSCSINHPIPFENSEVITNITFDVDSKASLGKLLKLLKANVTSENMPTNKTEF 900
DB 855 SQGTWSTSCSINHPIFRGGAQITFLATFDVSPKRAVLGDRLLLTANVSSENTPRTSKTTF 914

QY 901 QLELPVKYAVYVMVTVSHGVSTKYLNFAS-ENTSRVMQHOYQVSNLQORSIPISLVLVP 959
DB 915 QLELPVKYAVYVMVTVSSHEQTKYLNFSSESEKESHVAMHRYQVNNLQORDLPVSNFWVP 974

QY 960 VRLNQTVIMDRPQVTFSENLSSSTCHTERLPSSHSDFLAELRKAPVNVNCSIAVCORICDI 1019
DB 975 VELNQEAVMWDMVEVSHPNESLRCSEKIAPPASDFLAHQKPNVLDCSIAGCLFRCDV 1034

QY 1020 PFGIQEENFATLKGNSLSPDWYIKTSHNHLIIVSTABILFNSDVSFTLLPGGAFVRQTE 1079
DB 1035 PFSVQSELDFTLKGNSLFGWVRQILQKVSVSVABITFDSYVSQLPQGEAFMRAQT 1094

QY 1080 TKVEPFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKRYKDMSE 1128
DB 1095 TVLEKYVHNPTPLIVGSSIGLLLLALITAVLYKVGFFKRYKEMMBE 1143

RESULT 12
ADQ17510
ID ADQ17510 standard; protein; 1163 AA.
XX
AC ADQ17510;
XX
DI 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated protein - SEQ ID 327.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 327; 210pp; English.
XX
CC The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC protein of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 1163 AA;
```

Query Match		59.1%; Score 3473; DB 8; Length 1163;
Best Local Similarity		61.3%; Pred. No. 6.9e-280;
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;		
Qy	1	FNLDTENAMTFQENARGFQGVVQLQGSFVVGAPQEIIVAAQNRSLVQCDYSTGSCBPI 60
Db	20	FNLDTEELTAFRVDSAGFQGVVQVANSVWVVGAPQKIAAQTGGLYQCYSTGACEPI 79
Qy	61	RLOVPVEAVNMSLGLSLAATTSFPLLACGPTVHQTCSNTYVVKGLCFGLFSGNLQQPOK 120
Db	80	GLQVPPEAVNMSLGLSLATTSFPLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR 137
Qy	121	FPEALRGCPQEDSDTAFLIDGSGSIIPHDFRMKEFVSTVMEQLKXSKTLFSLMQYSEEF 180
Db	138	LPVSRQECFQEQDIFVLIDGSGSISSRNFAFMNFRAVISQFORPSTQFSLMQFSNKF 197
Qy	181	RIHFTFKFQNNPRSLVKPITQLGRTHATGVKRVIRELLNITNGARKNAFKILLIVI 240
Db	198	QTHFTFEFRITSNPLSLASVHQLQGYTYTATAIQNVVHRLFHASYGARRDATKILIVI 257
Qy	241	TGGEKFGDPLGYEDVIEADREGVIRYVIGVDAPRSEKSRQELNTIASKPPRDHVFQVN 300
Db	258	TGKKEGDSLDYKDVIPWADAAGIIRYVIGLAFQNRNSWKELNDIASQSEHIFKVE 317
Qy	301	NFEALKTIQNLREKIFAIEGTQTGSSSPHEMSQEGFSAAITNSGPELLSTVGSYDWAG 360
Db	318	DFDALKIQNLREKIFAIEGTQTGSSSPHEMSQEGFSAVFTPDGVLGAVGFTWSG 377
Qy	361	GVFLYTSKESKSTFLNMTRVSDMDAVLGYAAAIILNRVOSLVIGAPRYOHIGLVAMFR 420
Db	378	GAFLYPPNMSPTFLNMSQENVDMRDSVLGYSTELALMKGVQSLVIGAPRYOHTKAVIFT 437
Qy	421	QNTQWMSNANVKGTQIGFYAGSLCSVDVDSNGSTDVLVIGAPHYVBOTRGGQVSCPL 480
Db	438	QVSRQWRMKAETVGTQIGFYAGSLCSVDVDTGSTDVLVIGAPHYVBOTRGGQVSCPL 497
Qy	481	PRGQARWQDAVLYGQGPQWGRFGAALTVLGVDNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db	498	PRGWR-RWDCDAVLYGQGPQWGRFGAALTVLGVDNGDKLTDVVI GAPGEENRGAVYLF 556
Qy	541	HGTSGSGISPSHQSRIAGSKLSPLOYFGQSLSGQDLTWGDLVLTGVAGHVLRLRSQ 600
Db	557	HGVLGFSISPSHQSRIAGSKLSPLOYFGQSLSGQDLTWGDLVLTGVAGHVLRLRSQ 616
Qy	601	PVLVRKAIMEFNPREVARNVFECDQVWKGKEAGEVRVCLHVOKSTEDRLREGIOISVVT 660
Db	617	PVLWVGVMQFIPAEIPRSFAFCEQVVSQTLVQSNICLYIDKRSKNLLGSDLOSST 676
Qy	661	YDLALDSGRPHSAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDVPSPVILRLNF 720
Db	677	LDLALDPGRLSPRATFQETKNRSLSRVVLGKAHCENFNLLPSCVEDSVTPITLRLNF 736
Qy	721	SLVGTPLSAFNLRPVLAEDAQRLETFALPFENKCGNDNICODDLITFSFMSLDCLVVG 780
Db	737	TLVGKPELLAFNLRLPMLAALQAFYFASLPFENKCGADHICQDNLGISFSFPGKSLVG 796
Qy	781	GPREFNVTVVRNDGEDSYRTQTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 840
Db	797	SNLELNAEVMWVNDGEDSYRTQTFPPDLDSYRKVSTLQNRQSRQSWRLACESASSTEV 854
Qy	841	SGALKSTSCSINHPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF 900
Db	855	SOGTWSTSCRINHLIFPGGAQITFLATFDVSPKAVLGDRLLLTANVSSENNTPRTSKTFF 914
Qy	901	QLELPVKYAVMYVTSVSHGVSTKYNLTAS-ENTSRVQHQYQVNSLQGRSLPISLVFLVP 959
Db	915	QLELPVKYAVMYVTSVSHGVSTKYNLTAS-ENTSRVQHQYQVNSLQGRSLPISLVFLVP 974
Qy	960	VRNLQTVIWDPRQVTFSENLSTCTHCKERLPSHSDFLAELRKPVYVNCISAVCQRLQCDI 1019
Db	975	VELNQAEMVMDVEVSHVPQNPRLCSSEKAPASDFLAHIQKNPVLDDCSIAQCLFRCDV 1034
Qy	1020	PFQIGQEEFNATLKNLSFDWYIKTSHNHLIVSTAEILLFNDVSFTLLPGQGAFRVSQTE 1079
Db	1035	PSFSVQEBLDFTLKGNLSFGWVRQLQKKVSVVAEITFTDTSVYSLPQGAFAFMAQTT 1094
Qy	1080	TKVEPFEVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKROYKQDMWSE 1128
Db	1095	TVLEKYKVHNTPLIVGSSIGGLLLALITAVLYKVGFFKROYKEMMEE 1143
RESULT 13		
ID	AA07120	
AC	AA07120 standard; protein; 1163 AA.	
XX	AA07120;	
XX	25-MAR-2003 (revised)	
DT	05-FEB-1991 (first entry)	
DE	p150.95 alpha subunit encoded by clone lambda47.	
XX	p150.95 leucocyte adhesion receptor alpha-subunit; hairy cell leukaemia; rhinovirus.	
KW	Synthetic.	
OS		
XX	Key	Location/Qualifiers
XX	Region	1..19
FT	Region	/label= signal peptide
FT	Region	20..44
FT	Modified-site	/label= N-terminus
FT	Modified-site	61..63
FT	Modified-site	/label= glycosylation site
FT	Modified-site	89..91
FT	Modified-site	/label= glycosylation site
FT	Modified-site	385..387
FT	Modified-site	/label= glycosylation site
FT	Modified-site	392..394
FT	Modified-site	/label= glycosylation site
FT	Modified-site	697..699
FT	Modified-site	/label= glycosylation site
FT	Modified-site	735..737
FT	Modified-site	/label= glycosylation site
FT	Modified-site	899..901
FT	Modified-site	/label= glycosylation site
FT	Modified-site	904..906
FT	Modified-site	/label= glycosylation site
FT	Modified-site	939..941
FT	Modified-site	/label= glycosylation site
FT	Modified-site	1050..1052
FT	Modified-site	/label= glycosylation site
FT	Domain	1108..1133
FT		/label= transmembrane
XX	WO9010646-A.	
PN	20-SEP-1990.	
XX	09-MAR-1990;	90WO-US001257.
XX	09-MAR-1990;	90WO-US001257.
XX	(DAND ) DANA FARBER CANCER INST INC.	
PI	Corbi AA, Springer TA;	
XX	WPI; 1990-304985/40.	
DR	N-PSDB; AAQ06068.	
XX	Treatment of viral esp. rhino-viral infection - by admin. of alpha sub-unit of p150.95 cell surface adhesion receptor, opt. together with a beta chain of CD-18 family.	
PS	Disclosure; Fig 3; 59pp; English.	



CC Clone lambda X47 was isolated from a cDNA library constructed from total  
CC RNA extracted from phorbol myristate acetate stimulated HL-60  
CC myelomonocytic cells. The library was screened with oligonucleotide  
CC probes based on tryptic peptide fragments of p150.95. The sequence can be  
CC attached to appropriate control elements and expressed in prokaryotic and  
CC eukaryotic cells. The protein can be used to treat or prevent rhinoviral  
CC infection because it interacts with ICM-1 and inhibits cell-virus  
CC attachment. It can also be used as an anti-inflammatory agent. See also  
CC AAQ06063-4, AAQ06068, AAR07121-8 and AAR07152-6. (Updated on 25-MAR-2003  
CC to correct PR field.) (Updated on 25-MAR-2003 to correct PA field.)  
XX  
XX  
SQ Sequence 1163 AA;  
  
Query Match 59.0%; Score 3464; DB 2; Length 1163;  
Best Local Similarity 61.1%; Pred. No. 3.9e-295;  
Matches 690; Conservative 138; Mismatches 279; Indels 6; Gaps 4;  
  
QY 1 FNLDTENAMTFOENARGFGQSVVOLQGSRRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60  
DB 20 FNLDETELTA PRVDSAGDGSVVQVANSVVVVGAPQKITAANQTGLYQCYSTGACEPI 79  
  
QY 61 RLQVPEAVNMSSGLSLAATTSPQLLACGPTVHTQCSNTYVKGCLFLFGSNLRQDPQK 120  
DB 80 GLQVPPPEAVNMSSGLSLAATTSPQLLACGPTVHTQCSNTYVKGCLFLFGSNLRQDPQK 137  
  
QY 121 FPEALRGCPQEDSDIAFLIDSGSII PHDFRMKEFVSTVMEQLKSKTILSLMQYSEEF 180  
DB 138 LPVSRQECPRQEDIVFLIDSGSISRRNFATMNFVRAVISQFQSTQSLMWFQSNKF 197  
  
QY 181 RIHTEKFEQNNPRSLVKPTQLLGRTHATGVRKVI RELLNITNGARKNAKILIVI 240  
DB 198 QTHFTFEFRNTSNPLSLASVHQLOGTYTATQNVVHRLPHASYGARRDATKILIVI 257  
  
QY 241 TDGERFGDPLGYEDVPEADREGVIRYVIGVGDAFRSEKSRQELNTIASKPRDRHVFQVN 300  
DB 258 TDGKEGDSLDYKDVIPWADAAGIIRYALGVGLAFQNRNSWKELNDAISKPSQEHIFKVE 317  
  
QY 301 NFEALKTQNLREKIFAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSVSGSYDNAG 360  
DB 318 DFDALDKIQNLKEKIFAIEGTQTGSSSFEHMSQEGFSAAITNSGPLLSVSGSYDNAG 377  
  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAIIILNRVQSLVLGAPRYQHIGLVAMFR 420  
DB 378 GAFLYPPNMSPTFINMSQENVDRDSYLGSTELALWKGVQSLVLGAPRYQHTGRAVIFT 437  
  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 480  
DB 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGQVSVCP 497  
  
QY 481 PRGQRARWOCDAVLIGEQQQWGRFGAALTVLGVNGDKLTDVATGAPGEENRGAVLYF 540  
DB 498 PRGWR-RMWCDVLYGEQHPWGRFGAALTVLGVNGDKLTDVATGAPGEENRGAVLYF 556  
  
QY 541 HGTSGSGISPSHSQRTAGSKLSPLOYFGQSLSGGQDLMGDLVLTGVAQGHVLLRSQ 600  
DB 557 HGVLPSPISPSHSQRTAGSKLSPLOYFGQSLSGGQDLMGDLVLTGVAQGHVLLRSQ 616  
  
QY 601 PVLVRKAIIMEFPREARVNECNDQVKGKEAGEVRVCLHVQKSTRDLREGQIOSVVT 660  
DB 617 PVLWVGVSQFIPAEIPRSFAFECREQVSEQTUVQSNICLYIDKRSKLLGSRDLQSSVT 676  
  
QY 661 YDLALDSGPHRASVNETKNTSTRTOVLGLTQTCETLKLQPNCEIDPVSIVLRNLF 720  
DB 677 LDIALDPGLSPRATFQETKNRSLSRVRLGKAHCENFNLLLPSCVEDSVTPITURLNF 736  
  
QY 721 SLVGTPLSAFNLPRVLAEDAQRLETFALPFPEKNCNDNICQDQLSITFSFMSLDLVVG 780  
DB 737 TLVKGKLLAFNRURPMLAALAQRYFTASLPFEKNCADHICQDNLGISTFPGKSLVVG 796  
  
QY 781 GPREFNVTVTRNDGEDSYRTQVTFPPFLDLSYKRVSTLQNRQSRWSRLACASSTEV 840  
DB 797 SNLELNAEVVMNDGEDSYGTITTFSHPAGLSYRYVAEGQKQGLRSLHLTCDSA--PVA 854

QY 841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANTSENMMPTNKTEF 900  
DB 855 SQGTWSTSCRINHLIFRGGAITFLATFDVSPRAVLGDRLLLTANVSSENNTPTSTKTF 914  
  
QY 901 QLELPKYAVYVMTSHGVSTKYLNTAS-ENTSRVMOHOYQVSNLQORSIPISLVLVP 959  
DB 915 QLELPKYAVYVTVSSHEQTKYLNFSSEBESHVAMHRYQVNNUGQRLDPVSNFWVP 974  
  
QY 960 VRLNQTVMRPPQVTFSENLSSTCHTKERLPSSDFLAELRKAPVNVCSIAVCQRIQCDI 1019  
DB 975 VELNQEAVMWDMVEVSHQPNESLRCSSQKIAPPASDFLAHTKQNVLDCSIAGCLRRFCDV 1034  
  
QY 1020 PFGIQIEEFNATLKNLSFDWYIKTSHNLLIYSTAEILNDSVFTLLPGQGAIVRSQTE 1079  
DB 1035 PFSFVQEEELFTLLKGNLSPGWVRQILQKVSVSVABITFTSVYSQLPQGERFMRAQT 1094  
  
QY 1080 TKVEPPEVNPPLPIVCGSSVGGILLALITAAALYKLGFEKRYKDMSE 1128  
DB 1095 TVLEKYKVHNPPTPLIVGSSIGALLLALITAVLYKVGFFKRYKEMEE 1143  
  
RESULT 14  
AAW65091  
ID AAW65091 standard; protein; 1163 AA.  
XX  
AC AAW65091;  
XX  
DT 28-SEP-1998 (first entry)  
XX  
DE Human Beta-integrin CD11c subunit protein.  
XX  
KW Beta-integrin alpha-d subunit; human; modulator; treatment; psoriasis;  
KW type-I diabetes; atherosclerosis; multiple sclerosis; asthma;  
KW lung inflammation; acute respiratory distress syndrome; CD11c subunit;  
KW rheumatoid arthritis.  
XX  
OS Homo sapiens.  
XX  
PN US5728533-A.  
XX  
PD 17-MAR-1998.  
XX  
PF 07-JUN-1995; 95US-00485618.  
XX  
PR 23-DEC-1993; 93US-00173497.  
PR 05-AUG-1994; 94US-00286889.  
PR 21-DEC-1994; 94US-00362652.  
XX  
PA (ICOS-) ICOS CORP.  
XX  
PI Van Der Vieren M, Gallatin WM;  
XX  
DR WPI; 1998-206565/18.  
XX  
PT Screening assay for modulators of integrin binding - using immobilised or  
PT labelled alpha-d polypeptide, useful for, e.g. treating type-I diabetes.  
XX  
PS Example 5; Fig 1A-D; 106pp; English.  
XX  
CC This sequence represents a human beta-integrin CD11c subunit which is  
CC used to describe a method for identifying compounds that modulate the  
CC interaction of the beta-integrin alpha-d subunit with a binding partner  
CC of alpha-d which involves contacting an alpha-d polypeptide with an alpha  
CC -d binding partner, one of which is immobilised and the other of which is  
CC labelled, in the presence of a test compound, and determining if the  
CC compound affects binding between the alpha-d polypeptide and alpha-d  
CC binding partner, where the alpha-d polypeptide is alpha-d or its fragment  
CC comprising the cytoplasmic, transmembrane or extracellular domain of  
CC alpha-d. Compounds that modulate alpha-d binding could be used to treat  
CC diseases such as type-I diabetes, atherosclerosis, multiple sclerosis,  
CC asthma, psoriasis, lung inflammation, acute respiratory distress syndrome  
CC and rheumatoid arthritis  
XX

SQ Sequence 1163 AA;

Query Match 58.7%; Score 3450; DB 2; Length 1163;  
Best Local Similarity 61.0%; Pred. No. 5.7e-278;  
Matches 689; Conservative 144; Microbial 200;

[illegible]

Qy	1020	PFGCIQEEFNATUKGNLSPDWYIKTSHNLLIVSTAEILFNDSVFTLLPGQAFVRSQTE	1079
Db	1035	PSFSQBELFTIIKGNLSFGWNRQIIQKVSVVSVAEIIIDFSTSVCGLPQGEAFMBAQTI	1094
Qy	1080	TKUEPEVPNPDLPIVGSSVGGILLALITAALYKLGFVKROYKNMMSE	1128
Db	1095	TVLKRYKHNPDIPLVGSSIGGILLALITAVLYKVGFFKROYKENMEE	1143

RESULT 15  
AAB07361  
ID AAB07361 standard; protein; 1163 AA.

AC AAB07361;

DT 17-JAN-2001 (first entry)

Human CD11c protein sequence

XX Human: macronhase infiltration

KW leukocyte integrin; Leu-CAM; leukointegrin; immune response;  
KW inflammation; leukocyte adhesion deficiency; LAD; Type I diabetes;  
KW atherosclerosis; multiple sclerosis; asthma;  
KW lung inflammation; acute respiratory distress syndrome; Crohn's disease;  
KW rheumatoid arthritis; central nervous system injury; CD11c

Homo sapiens

PN WO200029446-A1.

25-MAY-2000

XX  
16 NOV 1988

Year	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	

PR 08-JUL-1999; 99US-00350259.

PA (ICOS-) ICOS CORP.

XX  
PI Gallatin MW Van Der Wier M.

XX  
WPT. 2000 387777/22

XX  
E  
C

PT macrophage infiltration and reduce inflammation at central nervous system injury sites.

Example 5: Fig 1: 270pp: English

Integrins are a class of membrane-associated molecules that participate in cellular adhesion. Integrins are made up of an alpha subunit and a beta subunit. One class of human integrins are restricted to expression in white blood cells and have a common beta2 subunit: the leukocyte integrins, L $\alpha$ -CAMs, leukointegrins or beta2 integrins. Beta2 integrins have an important role in immune and inflammatory responses. The present protein sequence is the human integrin alpha subunit C11c. This sequence was used in an alignment to identify a novel beta2 integrin alpha subunit: alpha d (AA66004 and AAB07359). The present sequence has approximately 66% identity to the protein sequence of alpha d. The Alpha d gene and protein may be useful in therapy for diseases linked to aberrant alpha d function e.g. Type I diabetes, atherosclerosis, multiple sclerosis, asthma, psoriasis, lung inflammation, acute respiratory distress syndrome, rheumatoid arthritis and leukocyte adhesion deficiency (LAD). In addition, anti-alpha d monoclonal antibodies may be used in the inhibition of macrophage infiltration at the site of a central nervous system injury. The monoclonal antibodies can also be used to detect and diagnose Crohn's disease.

Sequence 1163 AA:

### Query Match

Query Match 58.7%; Score 3450; DB 3; Length 1163;  
Best Local Similarity 61.0%; pred. No. 5.7e-278;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 12:17:49 ; Search time 97.5 Seconds  
(without alignments)  
4119.157 Million cell updates/sec

Title: US-09-902-481B-6

Perfect score: 5875

Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKRQYKDMSEGPPGAEPQ 1137

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1566620 seqs, 353225886 residues

Total number of hits satisfying chosen parameters: 1566620

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubaa/FCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5875	100.0	1137	10	US-09-902-481A-6
2	5868	99.9	1153	9	US-09-350-259-3
3	5868	99.9	1153	9	US-09-902-481A-1
4	5868	99.9	1153	14	US-09-891-943-3
5	5868	99.9	1153	14	US-10-144-259-30
6	5868	99.9	1153	14	US-10-207-655-176
7	5855	99.7	1137	10	US-09-902-481A-5
8	5852.5	99.6	1152	9	US-09-945-265-4
9	5845	99.5	1137	10	US-09-902-481A-4
10	5832	99.3	1137	10	US-09-902-481A-3
11	3473	59.1	1163	14	US-10-116-275-204
12	3450	58.7	1163	9	US-09-350-259-4
13	3450	58.7	1163	10	US-09-891-943-4

14	3417	58.2	1161	9	US-09-350-259-2
15	3417	58.2	1161	10	US-09-891-943-2
16	3401.5	57.9	1161	9	US-09-350-259-99
17	3401.5	57.9	1161	10	US-09-891-943-99
18	3239.5	55.1	1161	9	US-09-350-259-55
19	3239.5	55.1	1161	10	US-09-891-943-55
20	3228	54.9	1161	9	US-09-350-259-53
21	3228	54.9	1161	10	US-09-891-943-53
22	3218.5	54.8	1151	9	US-09-350-259-37
23	3218.5	54.8	1151	10	US-09-891-943-37
24	3207	54.6	1155	9	US-09-350-259-46
25	3207	54.6	1155	10	US-09-891-943-46
26	1848	31.5	369	13	US-10-087-192-1212
27	1552.5	26.4	1170	15	US-10-261-164-1
28	1548.5	26.4	1223	16	US-10-408-765A-295
29	1525	26.0	1086	16	US-10-408-765A-1871
30	1350.5	23.0	494	9	US-09-350-259-103
31	1229.5	20.9	494	10	US-09-891-943-103
32	1229.5	20.9	413	9	US-09-350-259-101
33	1157.5	19.7	413	10	US-09-891-943-101
34	1157.5	19.7	1179	14	US-10-177-550-2
35	1149	19.6	1179	14	US-10-173-551-2
36	1149	19.6	1179	14	US-09-984-130-103
37	1102.5	18.8	1151	10	US-09-836-353A-103
38	1102.5	18.8	1179	10	US-09-918-715-250
39	1102.5	18.8	1179	17	US-10-474-794-250
40	1102.5	18.8	1188	14	US-10-291-265-810
41	1101	18.7	1188	14	US-10-291-265-810
42	1097	18.7	1188	14	US-10-291-265-338
43	1093.5	18.6	1189	10	US-09-984-130-35
44	1093.5	18.6	1189	10	US-09-836-353A-35
45	1093.5	18.6	1189	15	US-10-262-839-4

#### ALIGNMENTS

##### RESULT 1

US-09-902-481A-6  
; Sequence 6, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RT/MS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-6

Query Match	100.0%	Score	5875;	DB	10;	Length	1137;
Best Local Similarity	100.0%	Pred. No.	0;	Mismatches	1137;	Conservative	0;
						Indels	0;
						Gaps	0;
QY	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYCCDYSTGSCPEI	60				
Db	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANQSGSLYCCDYSTGSCPEI	60				
QY	61	RLOVPVEAVNMSLGLSLAATTPSPOLLACGPTVHQTCSNTYVVKGLCFELFSGNLRRQQPQK	120				
Db	61	RLOVPVEAVNMSLGLSLAATTPSPOLLACGPTVHQTCSNTYVVKGLCFELFSGNLRRQQPQK	120				

Tue Nov 9 12:56:12 2004

Patent No. US20020062008A1

GENERAL INFORMATION:  
APPLICANT: Gallatin, Michael W.  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. US20020062008A1el Human 2

FILE REFERENCE: 27866/35004  
CURRENT APPLICATION NUMBER: US/09/350,259  
CURRENT FILING DATE: 1999-07-08  
EARLIER APPLICATION NUMBER: 09/193,043  
EARLIER FILING DATE: 1998-11-16  
EARLIER APPLICATION NUMBER: 08/173,497  
EARLIER FILING DATE: 1993-12-23  
EARLIER APPLICATION NUMBER: 08/286,889  
EARLIER FILING DATE: 1994-08-05  
EARLIER APPLICATION NUMBER: 08/362,652  
EARLIER FILING DATE: 1994-12-21  
EARLIER APPLICATION NUMBER: 08/943,363  
EARLIER FILING DATE: 1997-10-03  
NUMBER OF SEQ ID NOS: 114  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 3

LENGTH: 1153  
TYPE: prf  
ORGANISM: Homo sapiens  
US-09-350-259-3

Query Match 99.9%; Score 5868; DB 9; Length 1153;

Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQIKSKTKLFSLMQYSEEF	180
Db	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKEFVSTVMEQIKSKTKLFSLMQYSEEF	180
QY	181	RIHFTFKFQNNPNRSLVKPIITQLGRTHATGVRKVIRELNITNGARKNAFKILIVI	240
Db	181	RIHFTFKFQNNPNRSLVKPIITQLGRTHATGVRKVIRELNITNGARKNAFKILIVI	240
QY	241	TDGEKFGDPLGYEDVPEADREGVIRVIVGVGDAFRSEKSRQELNTIASPPDRHVFQVN	300
Db	241	TDGEKFGDPLGYEDVPEADREGVIRVIVGVGDAFRSEKSRQELNTIASPPDRHVFQVN	300
QY	301	NFEALKTIONQLEKIFAIEGTQTGSSSSPEHEMSQEGFSAATTSNGPLLTSTVGSYDWAG	360
Db	301	NFEALKTIONQLEKIFAIEGTQTGSSSSPEHEMSQEGFSAATTSNGPLLTSTVGSYDWAG	360
QY	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR	420
Db	361	GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRVQSLVLGAPRYOHIGLVAMFR	420
QY	421	QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL	480
Db	421	QNTGMWESNANVKGTQICAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL	480
QY	481	PRGQARWQCDAVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF	540
Db	481	PRGQARWQCDAVLYGEGQPGWRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAVYLF	540
QY	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ	600
Db	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTVGAQGHVLLRSQ	600
QY	601	PVLRVKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
Db	601	PVLRVKATMEFNPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT	660
QY	661	YDLALDSGRPHSAFVNETKSTRQTQVLGHTQTCETIKLOLPNCIEDPVSPVILRLNF	720
Db	661	YDLALDSGRPHSAFVNETKSTRQTQVLGHTQTCETIKLOLPNCIEDPVSPVILRLNF	720
QY	721	SLVGTPLSAFGNLRPVLAEDAQLFTALPFPEKNCNDNICODDLITFSFMSLDCLVVG	780
Db	721	SLVGTPLSAFGNLRPVLAEDAQLFTALPFPEKNCNDNICODDLITFSFMSLDCLVVG	780
QY	781	GPREFNVTVTRNDGDSYRQTQVTFPPDLDSYRKVSTLQNRQSRSWRLACESASTEV	840
Db	781	GPREFNVTVTRNDGDSYRQTQVTFPPDLDSYRKVSTLQNRQSRSWRLACESASTEV	840
QY	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF	900
Db	841	SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRTNKTEF	900
QY	901	QLELPVKIAYVMVTSHGVS TKYLNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV	960
Db	901	QLELPVKIAYVMVTSHGVS TKYLNFTASENTSRVMQHOYQVSNLQORSIPISLVFLVPV	960
QY	961	RLMQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP	1020
Db	961	RLMQTVLWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCDIP	1020
QY	1021	PFQIQEENATLKNLSFDMYIKTSHNHLITVSTAEILFNDVSFTLLPGQAFVRSQDET	1080
Db	1021	PFQIQEENATLKNLSFDMYIKTSHNHLITVSTAEILFNDVSFTLLPGQAFVRSQDET	1080
QY	1081	KVEPFEVNPPLIVGSSVGLLALITALYKLGFFPKROYKDMMSGGPPGAEPPQ	1137
Db	1081	KVEPFEVNPPLIVGSSVGLLALITALYKLGFFPKROYKDMMSGGPPGAEPPQ	1137

QY 661 YDLALDSGRPHRAVENETKNSRRQTVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720  
DB 677 YDLALDSGRPHRAVENETKNSRRQTVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 736  
QY 721 SLVGTPLSAFGLNRPVLADDAQLFTALPFEKNCNGNDNICDDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFGLNRPVLADDAQLFTALPFEKNCNGNDNICDDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPFDLDSYRKVSTLQNSORSQSWELACESASSTEV 840  
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPFDLDSYRKVSTLQNSORSQSWELACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVMVVTSYKYLNFTASENTSRVMQHQYQVSNLQORSQSWELACESASSTEV 960  
DB 917 QLELPVKYAVMVVTSYKYLNFTASENTSRVMQHQYQVSNLQORSQSWELACESASSTEV 976  
QY 961 RLNTQVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCIAVCORIQCDDIP 1020  
DB 977 RLNTQVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCIAVCORIQCDDIP 1036  
QY 1021 FFGIOEENATLKGNSLDFWIKTSHNHLIVSTAEILLFNDVSFTLLPGQAGFVRSQTEP 1080  
DB 1037 FFGIOEENATLKGNSLDFWIKTSHNHLIVSTAEILLFNDVSFTLLPGQAGFVRSQTEP 1096  
QY 1081 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKROYKDMSEGGPPGAEQ 1137  
DB 1097 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKROYKDMSEGGPPGAEQ 1153

## RESULT 3

US-09-902-481A-1  
; Sequence 1, Application US/09902481A  
; Publication No. US20030054440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RPT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902, 481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: mat\_peptide  
; LOCATION: (17)..()  
; OTHER INFORMATION:  
US-09-902-481A-1

Query Match  
Best Local Similarity 99.9%; Score 5868; DB 10; Length 1153;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLQGRVVVVGAPQEIIVAAANQORSGLYCCDYSTGSCBPI 60  
DB 17 FNLDTENAMTFOENARGFGQSVVQLQGRVVVVGAPQEIIVAAANQORSGLYCCDYSTGSCBPI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVVKGLCLFLFGSNLRQOPQK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPOLLACGPTVHTCSENTYVVKGLCLFLFGSNLRQOPQK 136

## RESULT 4

US-09-891-943-3  
; Sequence 3, Application US/09891943  
; Publication No. US2003007728A1

QY 121 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTVMEOLKKSKTLFSLMOYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLIDSGSIIIPHDFRMKEFVSTVMEOLKKSKTLFSLMOYSEEF 196  
QY 181 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIRELNITNGARKNAFKILIVI 240  
DB 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHATGVRKVIRELNITNGARKNAFKILIVI 256  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDARSEKSRQBELNITIAKPPDRHVFQVN 300  
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDARSEKSRQBELNITIAKPPDRHVFQVN 316  
QY 301 NFEALKTIONQUREKIPALEGTQTCSSSSFEHMSQEGFSAITNSGPELLSTVGSVDNAG 360  
DB 317 NFEALKTIONQUREKIPALEGTQTCSSSSFEHMSQEGFSAITNSGPELLSTVGSVDNAG 376  
QY 361 GVFLYTSKEKSTFINMTVRDSDMDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 420  
DB 377 GVFLYTSKEKSTFINMTVRDSDMDAYLGAAAAIILNRVQSLVLGAPRYQHIGLVAMFR 436  
QY 421 QNTGWESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLIGAPHYEQTGRGQVSVCP 480  
DB 437 QNTGWESNANVKGTOIGAYFGASICSVDVDSNGSTDLVLIGAPHYEQTGRGQVSVCP 496  
QY 481 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
DB 497 PRGORARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVGAGHVILLRSQ 600  
DB 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGOSLGGQDLTMDGLVDLTVGAGHVILLRSQ 616  
QY 601 PVLKVKALMEFNPREVARNVFECDQVVKKEAGEVRLVHVKQSTDRDLREGQIQSVVT 660  
DB 617 PVLKVKALMEFNPREVARNVFECDQVVKKEAGEVRLVHVKQSTDRDLREGQIQSVVT 676  
QY 661 YDLALDSGRPHRAVENETKNSRRQTVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 720  
DB 677 YDLALDSGRPHRAVENETKNSRRQTVGLTQTCETLKLQLPNCIEDPVSIVLRLNF 736  
QY 721 SLVGTPLSAFGLNRPVLADDAQLFTALPFEKNCNGNDNICDDLSITFSFMSLDCLVVG 780  
DB 737 SLVGTPLSAFGLNRPVLADDAQLFTALPFEKNCNGNDNICDDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPFDLDSYRKVSTLQNSORSQSWELACESASSTEV 840  
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPFDLDSYRKVSTLQNSORSQSWELACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900  
DB 857 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVMVVTSYKYLNFTASENTSRVMQHQYQVSNLQORSQSWELACESASSTEV 960  
DB 917 QLELPVKYAVMVVTSYKYLNFTASENTSRVMQHQYQVSNLQORSQSWELACESASSTEV 976  
QY 961 RLNTQVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCIAVCORIQCDDIP 1020  
DB 977 RLNTQVIWDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCIAVCORIQCDDIP 1036  
QY 1021 FFGIOEENATLKGNSLDFWIKTSHNHLIVSTAEILLFNDVSFTLLPGQAGFVRSQTEP 1080  
DB 1037 FFGIOEENATLKGNSLDFWIKTSHNHLIVSTAEILLFNDVSFTLLPGQAGFVRSQTEP 1096  
QY 1081 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKROYKDMSEGGPPGAEQ 1137  
DB 1097 KVEPFEVNPPLPLIVGSSVGLLLALITAAALYKLGFFKROYKDMSEGGPPGAEQ 1153

Tue Nov 9 12:56:12 2004

```

; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-891-943-3

Query Match 99.9%; Score 5868; DB 10; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLQGRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARFGQSVVQLQGRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 76

61 RLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFPGSNLRQOPQK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFPGSNLRQOPQK 136

121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSEEF 196

181 RIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGVRKVIKRELLNITNGARKNAFKILIVI 256

241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTIASKPRDHVFOYN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTIASKPRDHVFOYN 316

301 NFPAKTIQNLREKIFAIBGTQTGSSSFEHEMSQEGFSAATTSNGPILLSTVGSYDMAG 360
DB 317 NFPAKTIQNLREKIFAIBGTQTGSSSFEHEMSQEGFSAATTSNGPILLSTVGSYDMAG 376

361 GVFLYTSKEKSTFINTRVDSQMDNDAVLGYAAAILLRNVQSLVLGAPRYQHTGLVAMFR 420
DB 377 GVFLYTSKEKSTFINTRVDSQMDNDAVLGYAAAILLRNVQSLVLGAPRYQHTGLVAMFR 436

421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 480
DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 496

481 PRGORARWQCDVILYGGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAYLF 540
DB 497 PRGORARWQCDVILYGGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAYLF 556

541 HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTAVGAQHVLLRSQ 600
DB 557 HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTAVGAQHVLLRSQ 616

601 PVLVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 660
DB 617 PVLVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 676

661 YDLALDGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 720
DB 677 YDLALDGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 736

721 SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFKNCGNDMNCQDDLSITFSFMSLDCLVVG 780
DB 737 SLVGTPLSAFGLNLRPVLAEDAQRLEFALFPFKNCGNDMNCQDDLSITFSFMSLDCLVVG 796

781 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSVRKYSTLQNRQSRQSWRLACESASSTEV 840
DB 797 GPREFNVTVVRNDGEDSYRTQVTFPFLDLSVRKYSTLQNRQSRQSWRLACESASSTEV 856

841 SGALKSTSCSINHIPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
DB 857 SGALKSTSCSINHIPIPPENSEVFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 916

901 QLELPVKYAVYVMTVSHGVSTKYLNFASENTSRVMOHQVQVSNLQORSPLISLVFLVPV 960
DB 917 QLELPVKYAVYVMTVSHGVSTKYLNFASENTSRVMOHQVQVSNLQORSPLISLVFLVPV 976

961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020
DB 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036

1021 FFGTQBEFNATLKGNSLDFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQAFVRSQTE 1080
DB 1037 FFGTQBEFNATLKGNSLDFDWYIKTSHNHLIVSTABILFNDVSFTLLPGQAFVRSQTE 1096

1081 KVEPFEPVNPPLVLIVGSSVGGLLALLIITAALYKLGFPFKQYKDMSEGGPPGAEPQ 1137
DB 1097 KVEPFEPVNPPLVLIVGSSVGGLLALLIITAALYKLGFPFKQYKDMSEGGPPGAEPQ 1153

RESULT 5
US-10-144-259-30
; Sequence 30, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US/10/144,259
; PRIOR FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-30

Query Match 99.9%; Score 5868; DB 14; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARFGQSVVQLQGRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARFGQSVVQLQGRVVVGAPOEIVAAANQSGSLYQCDYSTGSCPEI 76

61 RLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFPGSNLRQOPQK 120
DB 77 RLQVPEAVNMSLGLSLAATTSPPQLLACGPTVHQTCTSENTYVKGCLFPGSNLRQOPQK 136

121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSEEF 180
DB 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSEEF 196

181 RIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGVRKVIKRELLNITNGARKNAFKILIVI 240
DB 197 RIHFTFKFQNNPNRSLVKPITOLLGRTHRTATGVRKVIKRELLNITNGARKNAFKILIVI 256

241 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTIASKPRDHVFOYN 300
DB 257 TDGEKFGDPLGYEDVPEADREGVIRYVIGVDGDAFSEKSRQELNTIASKPRDHVFOYN 316

301 NFPAKTIQNLREKIFAIBGTQTGSSSFEHEMSQEGFSAATTSNGPILLSTVGSYDMAG 360
DB 317 NFPAKTIQNLREKIFAIBGTQTGSSSFEHEMSQEGFSAATTSNGPILLSTVGSYDMAG 376

361 GVFLYTSKEKSTFINTRVDSQMDNDAVLGYAAAILLRNVQSLVLGAPRYQHTGLVAMFR 420
DB 377 GVFLYTSKEKSTFINTRVDSQMDNDAVLGYAAAILLRNVQSLVLGAPRYQHTGLVAMFR 436

421 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 480
DB 437 QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEQTRGGQVSCPL 496

481 PRGORARWQCDVILYGGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAYLF 540
DB 497 PRGORARWQCDVILYGGQGPWGRFGAALTIVLGDVNGDKLTDVAIGAPGEDNKGAYLF 556

541 HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTAVGAQHVLLRSQ 600
DB 557 HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTAVGAQHVLLRSQ 616

601 PVLVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 660
DB 617 PVLVKAIIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGIQSVVT 676

661 YDLALDGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 720
DB 677 YDLALDGRPHSRVFNETHKSTRQTVLGLTQTCETLKLQLPNCIEDVSPVILRLNF 736
```



Db 197 RIHFTFKFQNNPNRSLVKPIITOLLGRTHTATGIRKVVRELNITNGARKNAFKILVVI 256  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVIVIGDAFRSEKSRQELNITIASKPPRDHVFQVN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRVIVIGDAFRSEKSRQELNITIASKPPRDHVFQVN 316  
QY 301 NFEALKTIQNLREKIFAIEGTGTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTGTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 376  
QY 361 GVFLYTSKESKSTPNNTRVDSMDNDAYLGVAALILRNVRQSLVILGAPRYOHIGLVAMFR 420  
Db 377 GVFLYTSKESKSTPNNTRVDSMDNDAYLGVAALILRNVRQSLVILGAPRYOHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 496  
QY 481 PRGQARWOCDAVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWOCDAVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKSPRLOYFGQSLSGGODLTMDGLVDTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKSPRLOYFGQSLSGGODLTMDGLVDTVGAQGHVLLRSQ 616  
QY 601 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660  
Db 617 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLCLVVG 780  
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLCLVVG 796  
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSORSWRACASSTEV 840  
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSORSWRACASSTEV 856  
QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916  
QY 901 QLELPVKYAVYVVTSHGVSTKYLNTASNTSRVMQHOYOVSNLGRSLPISLVFLVPV 960  
Db 917 QLELPVKYAVYVVTSHGVSTKYLNTASNTSRVMQHOYOVSNLGRSLPISLVFLVPV 976  
QY 961 RLNOTVWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDIP 1020  
Db 977 RLNOTVWRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORICDIP 1036  
QY 1021 FFGIQEENATLKNLSFDWYIKTSHNLLIYVSTABILNDSVFTLLPQGAFAVSQET 1080  
Db 1037 FFGIQEENATLKNLSFDWYIKTSHNLLIYVSTABILNDSVFTLLPQGAFAVSQET 1096  
QY 1081 KVEPEVNPPLIYGVSSVGGILLALITAAALYKLGFFKRYOKDMSEGGPPGABPQ 1137  
Db 1097 KVEPEVNPPLIYGVSSVGGILLALITAAALYKLGFFKRYOKDMSEGGPPGABPQ 1153

## RESULT 6

US-10-207-655-176  
; Sequence 176, Application US/10207655  
; Publication No. US20030118592A1  
; GENERAL INFORMATION:  
; APPLICANT: Ledbetter, Jeffrey A.  
; APPLICANT: Hayden-Ledbetter, Martha S.  
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
; FILE REFERENCE: 390069, 401C1  
; CURRENT APPLICATION NUMBER: US/10/207,655

; CURRENT FILING DATE: 2002-07-25  
; NUMBER OF SEQ ID NOS: 426  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 176  
; LENGTH: 1153  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-207-655-176

Query Match 99.9%; Score 5868; DB 14; Length 1153;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVVGAPQIEIVAAANORGSLYQCDYSTGSCBPI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVVGAPQIEIVAAANORGSLYQCDYSTGSCBPI 76  
QY 61 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGLCFLFGSNLRQPOK 120  
Db 77 RLOVPEAVNMSLGLSLAATTSPOLLACGPTVHQTCSNTYVKGLCFLFGSNLRQPOK 136  
QY 121 PPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKFVSTVMEQLKSKTFLSLMQYSBEF 180  
Db 137 PPEALRGCPQEDSDIAFLIDGSGSIIPHDPRMKFVSTVMEQLKSKTFLSLMQYSBEF 196  
QY 181 RIHFTFKFQNNPNRSLVKPIITOLLGRTHTATGIRKVVRELNITNGARKNAFKILVVI 240  
Db 197 RIHFTFKFQNNPNRSLVKPIITOLLGRTHTATGIRKVVRELNITNGARKNAFKILVVI 256  
QY 241 TDGEKFGDPLGYEDVPEADREGVIRVIVIGDAFRSEKSRQELNITIASKPPRDHVFQVN 300  
Db 257 TDGEKFGDPLGYEDVPEADREGVIRVIVIGDAFRSEKSRQELNITIASKPPRDHVFQVN 316  
QY 301 NFEALKTIQNLREKIFAIEGTGTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTGTGSSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 376  
QY 361 GVFLYTSKESKSTPNNTRVDSMDNDAYLGVAALILRNVRQSLVILGAPRYOHIGLVAMFR 420  
Db 377 GVFLYTSKESKSTPNNTRVDSMDNDAYLGVAALILRNVRQSLVILGAPRYOHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 480  
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYEEOTRGQVSVCP 496  
QY 481 PRGQARWOCDAVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWOCDAVLYGEOQPGWRFGAALTIVLDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKSPRLOYFGQSLSGGODLTMDGLVDTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKSPRLOYFGQSLSGGODLTMDGLVDTVGAQGHVLLRSQ 616  
QY 601 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 660  
Db 617 PVLRVKAIMEFNPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 720  
Db 677 YDLALDSGRPHSRVAFNETKSTRQTOVLGLTQTCETLKLQLPNCIEDPVSPIVLRNF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLCLVVG 780  
Db 737 SLVGTPLSAFGLNRPVLAEDAQRFTALFPPEKNCNDNICODDLSITFSFMSLCLVVG 796  
QY 781 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSORSWRACASSTEV 840  
Db 797 GPREFNVTVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQVRSORSWRACASSTEV 856  
QY 841 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
Db 857 SGALKSTSCSINHPIFENSEVTNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916

Qy	901	OLELPVKYAVYVMTSHGVTSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLFLVLPV	960
Db	917	QLELPVKYAVYVMTSHGVTSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLFLVLPV	976
Qy	961	RLNQTWIDRPOVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP	1020
Db	977	RLNQTWIDRPOVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP	1036
Qy	1021	FFGIQEEFNATLKGNSFDWYIKTSHNHLIIIVSTAELFNDSVFTLLPGQGFVRSQTET	1080
Db	1037	FFGIQEEFNATLKGNSFDWYIKTSHNHLIIIVSTAELFNDSVFTLLPGQGFVRSQTET	1096
Qy	1081	KVEPFEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYQKMMSEGGPPGAEPO	1137
Db	1097	KVEPFEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYQKMMSEGGPPGAEPO	1153
RESULT 7			
US-09-902-481A-5			
; Sequence 5, Application US/09902481A			
; Publication No. US20030054440A1			
; GENERAL INFORMATION:			
; APPLICANT: Springer, Timothy A.			
; APPLICANT: Shimaoka, Motomu			
; APPLICANT: Shifman, Julia			
; APPLICANT: Mayo, Stephen			
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY			
; FILE REFERENCE: A-70586-1/RT/RMS/RMK			
; CURRENT APPLICATION NUMBER: US/09/902,481A			
; CURRENT FILING DATE: 2001-07-09			
; PRIOR APPLICATION NUMBER: US 60/216,600			
; PRIOR FILING DATE: 2000-07-07			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 5			
; LENGTH: 1137			
; TYPE: PRT			
; ORGANISM: Artificial sequence			
; FEATURE:			
; OTHER INFORMATION: synthetic			
US-09-902-481A-5			
Query Match 99.7%; Score 5855; DB 10; Length 1137;			
Best Local Similarity 99.1%; Pred. No. 0;			
Matches 1127; Conservative 8; Mismatches 2; Indels 0; Gaps 0;			
Qy	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAVNRGSLYQCDYSTGSCPEI	60
Db	1	FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPOEIVAAVNRGSLYQCDYSTGSCPEI	60
Qy	61	RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVYKGLCFPLGSLNRQOPQK	120
Db	61	RLQVPVEAVNMSLGLSLAATTSPPQLACGPTVHQTCSNTYVYKGLCFPLGSLNRQOPQK	120
Qy	121	FPEALRGCPEDSDIAPLDGSGSIIPDHPFRKKEFVSTVMEQLKSKTLFSLMOYSEEF	180
Db	121	FPEALRGCPEDSDIAPLDGSGSIIPDHPFRKKEFVSTVMEQLKSKTLFSLMOYSEEF	180
Qy	181	RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILIVI	240
Db	181	RIHFTFKFQNNPNRSLVKPIITOLLGRTHATGVRKVIKRELLNITNGARKNAFKILILI	240
Qy	241	TGDKFGDPLGYEDVIEADREGVIRYVIGVDAPFRSEKSKQELNTASKPRPHVQVN	300
Db	241	TGDKFGDPLGYEDVIEADREGVIRYVIGVDAPFRSEKSKQELNTASKPRPHVQVN	300
Qy	301	NFEALKTQNLKREKIFAEIGTQSGSSSFHEMSQEGFSAAITNSGFLLLTVGSYDAG	360
Db	301	NFEALKTQNLKREKIFAEIGTQSGSSSFHEMSQEGFSAAITNSGFLLLTVGSYDAG	360
Qy	361	GVFLYTSKESKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVILGAPRYQHIGLVAMFR	420
Db	361	GVFLYTSKESKSTFINMTVDSMDNDAYLGYAAAIILNRVQSLVILGAPRYQHIGLVAMFR	420

Qy	421	QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYIEOTRGQSVSVCPL	480
Db	421	QNTGMWESNANVGTQIGAYFGASLCSVDVDSNGSDTLVLIGAPHYIEOTRGQSVSVCPL	480
Qy	481	PRGORARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAAYILF	540
Db	481	PRGORARWQCDVLYGEOQPMGRFGAALTVLGDVNGDKLTDVAIGAPCEENRGAAYILF	540
Qy	541	HGTSGSGISPSHSORISAGSKLSPRLQYFGQSISGGQDLTMDGLVDLTVGAGQHVLLLRQ	600
Db	541	HGTSGSGISPSHSORISAGSKLSPRLQYFGQSISGGQDLTMDGLVDLTVGAGQHVLLLRQ	600
Qy	601	PVLURVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRREGIQSVVT	660
Db	601	PVLURVKAIMEFNPREVARNVFECDQVVKGEAGEVRVCLHVOKSTRDRREGIQSVVT	660
Qy	661	YDLALDSGRPHSRAVFNENKSTRQVGLGTQTCETLKLQPLNCIEDPVSPIVLRNF	720
Db	661	YDLALDSGRPHSRAVFNENKSTRQVGLGTQTCETLKLQPLNCIEDPVSPIVLRNF	720
Qy	721	SLVGTPLSAFGLNLRPVLAEDAQRLFTALFPFEKNCVGNNDICQDDLSITFSFMSLDCVLVG	780
Db	721	SLVGTPLSAFGLNLRPVLAEDAQRLFTALFPFEKNCVGNNDICQDDLSITFSFMSLDCVLVG	780
Qy	781	GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV	840
Db	781	GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSYRKVSTLQNRQORSWRLACESASSTEV	840
Qy	841	SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
Db	841	SGALKSTSCSINHPIFPENSEVTNITPDVDSKASLGNKLLKANVTSENMPRTNKTEF	900
Qy	901	OLELPVKYAVYVMTSHGVTSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLFLVLPV	960
Db	901	OLELPVKYAVYVMTSHGVTSTKYLNTASNTSRVMOHQYQVSNLQORSPLISLFLVLPV	960
Qy	961	RLNQTWIDRPOVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP	1020
Db	961	RLNQTWIDRPOVTFSENLSSCHTKERLPSHSDFLAELRKAPVNVNCSIAVCORIQCDIP	1020
Qy	1021	FFGIQEEFNATLKGNSFDWYIKTSHNHLIIIVSTAELFNDSVFTLLPGQGFVRSQTET	1080
Db	1021	FFGIQEEFNATLKGNSFDWYIKTSHNHLIIIVSTAELFNDSVFTLLPGQGFVRSQTET	1080
Qy	1081	KVEPFEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYQKMMSEGGPPGAEPO	1137
Db	1081	KVEPFEVNPPLPIVGVSSVGGLLALLALITAAALYKLGFFKQYQKMMSEGGPPGAEPO	1137
RESULT 8			
US-09-945-265-4			
; Sequence 4, Application US/09945265			
; Patent No. US20020123614A1			
; GENERAL INFORMATION:			
; APPLICANT: Springer, Timothy A.			
; APPLICANT: Shimaoka, Motomu			
; APPLICANT: Lu, Chafen			
; TITLE OF INVENTION: MODIFIED POLYPEPTIDES STABILIZED IN A			
; FILE REFERENCE: CBN-002CP			
; CURRENT APPLICATION NUMBER: US/09/945,265			
; CURRENT FILING DATE: 2001-08-31			
; PRIOR APPLICATION NUMBER: US 60/229,700			
; PRIOR FILING DATE: 2000-09-01			
; NUMBER OF SEQ ID NOS: 4			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 4			
; LENGTH: 1152			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-09-945-265-4			

```
Query Match          99.6%; Score 5852.5; DB 9; Length 1152;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1132; Conservative 3; Mismatches 1; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
DB 17 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
QY 61 RLOQVPEAVNMSLGLSLAATTPPOLLACGPTVHQTCSSENTYVKGCLCFLFGSNLROPOK 120
DB 77 RLOQVPEAVNMSLGLSLAATTPPOLLACGPTVHQTCSSENTYVKGCLCFLFGSNLROPOK 136
QY 121 FPEARLGGCQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQLKSKTLFSLMOYSEEF 180
DB 137 FPEARLGGCQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQLKSKTLFSLMOYSEEF 196
QY 181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
DB 197 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 256
QY 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAFRSEKSRQELNTIASKPPDRDHVFOVN 300
DB 257 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAFRSEKSRQELNTIASKPPDRDHVFOVN 316
QY 301 NFEALKTIONOLREKIFAIEGTQTGSSSFHEMSQEGFSAATNSGPELLSTVGSYDWAG 360
DB 317 NFEALKTIONOLREKIFAIEGTQTGSSSFHEMSQEGFSAATNSGPELLSTVGSYDWAG 376
QY 361 GVFLYTSKEKSTFINMTTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
DB 377 GVFLYTSKEKSTFINMTTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 436
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCPL 480
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCPL 496
QY 481 PRGQARWQCDVILYGCQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 497 PRGQARWQCDVILYGCQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 555
QY 541 HGTSGSGISPSHQSRIAGSKSLPRLQYFGOSLGGQDLTMDGLVDLTVGAGQHVLRLSQ 600
DB 556 HGTSGSGISPSHQSRIAGSKSLPRLQYFGOSLGGQDLTMDGLVDLTVGAGQHVLRLSQ 615
QY 601 PVLRVKAIEMFNPREVARNVFECDQVVGKGEAGEVRVCLHVOKSTRDLRREGIOQSVVT 660
DB 616 PVLRVKAIEMFNPREVARNVFECDQVVGKGEAGEVRVCLHVOKSTRDLRREGIOQSVVT 675
QY 661 YDLALDSGRPHRAVENETKNSTRQTQVILGTQTCETLKLQLPNCIEDPVSPVILRLNF 720
DB 676 YDLALDSGRPHRAVENETKNSTRQTQVILGTQTCETLKLQLPNCIEDPVSPVILRLNF 735
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLFTALPFFKNCNDNI CODDLSITFSFMSLCLVVG 780
DB 736 SLVGTPLSAFGLNRPVLAEDAQRLFTALPFFKNCNDNI CODDLSITFSFMSLCLVVG 795
QY 781 GREFNVTTVRNDEGDSYRTQVTFPPFLDLSYRKVSTIQNORSOKSWELACESASTEV 840
DB 796 GREFNVTTVRNDEGDSYRTQVTFPPFLDLSYRKVSTIQNORSOKSWELACESASTEV 855
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900
DB 856 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 915
QY 901 QLELPVKYAVVMVTSHGVSYKYLNFNTASENTSRVMQHQYQVSNLQORSLPISLVFLVPV 960
DB 916 QLELPVKYAVVMVTSHGVSYKYLNFNTASENTSRVMQHQYQVSNLQORSLPISLVFLVPV 975
QY 961 RLNQTVIWRDQVTFSENLSTCHTKERLPKSHDPLAELKAPVNVNCIAVCQRIQCDIP 1020
DB 976 RLNQTVIWRDQVTFSENLSTCHTKERLPKSHDPLAELKAPVNVNCIAVCQRIQCDIP 1035
QY 1021 FFGIQEENATLKGSLFDWYIKTSHNLLIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1080
DB 1036 FFGIQEENATLKGSLFDWYIKTSHNLLIVSTAEILFNDVSFTLLPGQCAFVRSOTET 1095
QY 1081 KVEPFEVPNPLPLIVGSSVGGLLLALITAAALYKLGFFKRYQKMMSEGGPPGAEPQ 1137
DB 1096 KVEPFEVPNPLPLIVGSSVGGLLLALITAAALYKLGFFKRYQKMMSEGGPPGAEPQ 1152

RESULT 9
US-09-902-481A-4
; Sequence 4, Application US/09902481A
; Publication No. US20030054440A1
; GENERAL INFORMATION:
; APPLICANT: Springer, Timothy
; APPLICANT: Shimaoka, Motomu
; APPLICANT: Shifman, Julia
; APPLICANT: Mayo, Stephen
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY
; FILE REFERENCE: A-70586-1/RFT/RMS/RMK
; CURRENT APPLICATION NUMBER: US/09/902,481A
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/216,600
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 4
; LENGTH: 1137
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: synthetic
US-09-902-481A-4

Query Match          99.5%; Score 5845; DB 10; Length 1137;
Best Local Similarity 99.0%; Pred. No. 0;
Matches 1126; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
DB 1 FNLDTENAMTFOENARGFGQSVVQLOGSRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60
QY 61 RLOQVPEAVNMSLGLSLAATTPPOLLACGPTVHQTCSSENTYVKGCLCFLFGSNLROPOK 120
DB 61 RLOQVPEAVNMSLGLSLAATTPPOLLACGPTVHQTCSSENTYVKGCLCFLFGSNLROPOK 120
QY 121 FPEARLGGCQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQLKSKTLFSLMOYSEEF 180
DB 121 FPEARLGGCQEDSDIAFLIDGSGSIIPHPFRMKFVSTVMEQLKSKTLFSLMOYSEEF 180
QY 181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
DB 181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI 240
QY 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAFRSEKSRQELNTIASKPPDRDHVFOVN 300
DB 241 TDGEKFGDPLGYEDVIPADREGVIRYVIGVDAFRSEKSRQELNTIASKPPDRDHVFOVN 300
QY 301 NFEALKTIONOLREKIFAIEGTQTGSSSFHEMSQEGFSAATNSGPELLSTVGSYDWAG 360
DB 301 NFEALKTIONOLREKIFAIEGTQTGSSSFHEMSQEGFSAATNSGPELLSTVGSYDWAG 360
QY 361 GVFLYTSKEKSTFINMTTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
DB 361 GVFLYTSKEKSTFINMTTRVDSMDNDAYLGAAAAIILNRVQSLVILGAPRYOHIGLVAMFR 420
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCPL 480
DB 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGGOVSVCPL 480
QY 481 PRGQARWQCDVILYGCQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
DB 481 PRGQARWQCDVILYGCQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
```

541 HGTSGGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
541 HGTSGGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
601 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDRRLREGQIQSVVT 660  
601 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDRRLREGQIQSVVT 660  
661 YDLALDSGRPHSRVAVNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSIVLRLNF 720  
661 YDLALDSGRPHSRVAVNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSIVLRLNF 720  
721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFEKNCNDNI CDDLSITFSFMSLDCLVVG 780  
721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFEKNCNDNI CDDLSITFSFMSLDCLVVG 780  
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKVKSTLQNSORSWRWLACESASSTEV 840  
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKVKSTLQNSORSWRWLACESASSTEV 840  
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900  
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900  
901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLSILVELVPV 960  
901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLSILVELVPV 960  
961 RLNQTVIWDROPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDDIP 1020  
961 RLNQTVIWDROPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDDIP 1020  
1021 FPGIOEEFNATLKGNSLFDWYIKTSHNLLIVSTAEILLFNDVSFTLLPGQAGFVRSQTET 1080  
1021 FPGIOEEFNATLKGNSLFDWYIKTSHNLLIVSTAEILLFNDVSFTLLPGQAGFVRSQTET 1080  
1081 KVEPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPO 1137  
1081 KVEPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPO 1137

RESULT 10  
US-09-902-481A-3  
; Sequence 3, Application US/09902481A  
; Publication No. US2003005440A1  
; GENERAL INFORMATION:  
; APPLICANT: Springer, Timothy  
; APPLICANT: Shimaoka, Motomu  
; APPLICANT: Shifman, Julia  
; APPLICANT: Mayo, Stephen  
; TITLE OF INVENTION: NOVEL PROTEINS WITH INTEGRIN-LIKE ACTIVITY  
; FILE REFERENCE: A-70586-1/RT/RMS/RMK  
; CURRENT APPLICATION NUMBER: US/09/902,481A  
; CURRENT FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: US 60/216,600  
; PRIOR FILING DATE: 2000-07-07  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1137  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: synthetic  
US-09-902-481A-3

Query Match 99.3%; Score 5832; DB 10; Length 1137;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 1122; Conservative 9; Mismatches 6; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQENARGFGQSVVQLQSGRRVVVGAPOETVAANQORGLSYQCDDYSTGSCBPI 60  
DB 1 FNLDTENAMTQENARGFGQSVVQLQSGRRVVVGAPOETVAANQORGLSYQCDDYSTGSCBPI 60

61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK 120  
61 RLQVPVEAVNMSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSNLRQPOK 120  
121 FPEALRGCEQSDIAFLIDGSGSIIIPHDPRMKBEFVSTVMEQLKSKSKTLFSLMQYSEEF 180  
121 FPEALRGCEQSDIAFLIDGSGSIIIPHDPRMKBEFVSTVMEQLKSKSKTLFSLMQYSEEF 180  
181 RHFTFKFQNNENPRLVKPIITQLLGRTHATGVRVKVIRELLNINANGARKNAFKILIVI 240  
181 RHFTFKFQNNENPRLVKPIITQLLGRTHATGVRVKVIRELLNINANGARKNAFKILIVI 240  
241 TDGEKFGDPLGYEDVPIPEADREGVIRYVIGVDAFSEKSEKSOELNTIASKPRDHVFOVN 300  
241 TDGEKFGDPLGYEDVPIPEADREGVIRYVIGVDAFSEKSEKSOELNTIASKPRDHVFOVN 300  
301 NFEALXTIQLNOLREKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG 360  
301 NFEALXTIQLNOLREKIFAIEGTOTGSSSSFEHMSQEGFSAATISNGPLLLSTVGSYDWAG 360  
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLYAAAIILNRVQSLVGLGAPRYQHIGLVAMER 420  
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLYAAAIILNRVQSLVGLGAPRYQHIGLVAMER 420  
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIAGPHYEOTRGGOVSVCP 480  
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIAGPHYEOTRGGOVSVCP 480  
481 PRGQARWQCDVAVLYGEGQPGWGFGAALTVLGVNGDKLTDVAITGAPGEEDNRGAVYLF 540  
481 PRGQARWQCDVAVLYGEGQPGWGFGAALTVLGVNGDKLTDVAITGAPGEEDNRGAVYLF 540  
541 HGTSGGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
541 HGTSGGISPSHSQRIAGSKLSPRLQYFGOSLGGQDLTMDGLVLDLTVGAQGHVLLRSQ 600  
601 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDRRLREGQIQSVVT 660  
601 PVLRVKAIMFNPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDRRLREGQIQSVVT 660  
661 YDLALDSGRPHSRVAVNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSIVLRLNF 720  
661 YDLALDSGRPHSRVAVNETKNSRTRQTVGLGTQTCETLKLQLPNCIEDPVSIVLRLNF 720  
721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFEKNCNDNI CDDLSITFSFMSLDCLVVG 780  
721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFEKNCNDNI CDDLSITFSFMSLDCLVVG 780  
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKVKSTLQNSORSWRWLACESASSTEV 840  
781 GPREFNVTVVRNDGEDSYRTQVTFPPDLVSRKVKSTLQNSORSWRWLACESASSTEV 840  
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900  
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENNMPRTNKTET 900  
901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLSILVELVPV 960  
901 QLELPVKYAVVMVVTSHGVSTKYLNFTASENTRVMQHOYQVSNLQORSLSILVELVPV 960  
961 RLNQTVIWDROPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDDIP 1020  
961 RLNQTVIWDROPQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCRIQCDDIP 1020  
1021 FPGIOEEFNATLKGNSLFDWYIKTSHNLLIVSTAEILLFNDVSFTLLPGQAGFVRSQTET 1080  
1021 FPGIOEEFNATLKGNSLFDWYIKTSHNLLIVSTAEILLFNDVSFTLLPGQAGFVRSQTET 1080  
1081 KVEPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPO 1137  
1081 KVEPFEVNPPLIVGSSVGGLLLALITAAALYKLGFFKRYKQKDMMSGGPPGAEPO 1137

RESULT 11  
US-10-116-275-204  
; Sequence 204, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 204  
; LENGTH: 1163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-275-204

Query Match 59.1%; Score 3473; DB 14; Length 1163;  
Best Local Similarity 61.3%; Pred. No. 6e-284;  
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARFGOSVVOQLQGSRRVVGAPQBIIVANORGSILYQCDYSGSCEPI 60  
DB 20 FNLDTTELTAFRVDSAGFSDVSVQYANSWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGFLFSGNLRQOPQK 120  
DB 80 GLQVPEAVNMSLGLSLASTTSQQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTOR 137  
QY 121 FPAALGCTQEDSDIAFLDGGSSIIIPDPRRMEKFVSTVMEQLKKSKTLFSLMOWSEEF 180  
DB 138 LPVSRQECRQEQDIFVLDDGSSISRRNPFATMNFVRAVISOFQPSQFSLMQFSNRF 197  
QY 181 RIHFTKEFQNNPNSLAKPTQLIGRTHATGVRKVIRELNIINGARKNAFKILIVI 240  
DB 198 QTHFTFEEPRTSNPLUSLASVHQLQGTYTATAIQNVVHRLFHASYGARRDATKILIVI 257  
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDHVFOYN 300  
DB 258 TDGKKGDSLDYKVTIPMADAGIIRYALGVGLAFQNRNSWKNELNDIAKPSQEHIFKYE 317  
QY 301 NFEALKTIONQUREKIFAIEGTQSSSFHEMSQEGFSAALITSNGPLLSVGSYDWAQ 360  
DB 318 DFDALDKIQNLQEKIFAIEGTETTSSSSFELEMAQEGFSAVFTPDGVLGAVGFTWSG 377  
QY 361 GVFLYTSKESFTFINNTRVDSMDNDAIYGAATILRNVSQSLVIGAPYQHIGLVAMFR 420  
DB 378 GAFLTPNMSPTFINMSQBNVMDRDSYLGTELALWKGVQSLVIGAPRYQHTGKAVIPT 437  
QY 421 QNTGMWESNANYKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQSVCP 480  
DB 438 QVSRQWRMKAETVGTQIGSYFGASLCSVDVDTGSDTDLVIGAPHYEQTRGGQSVCP 497  
QY 481 PRGQARWQCDVAVLGEQCPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRCVAVLF 540  
DB 498 PRGWR-RWMCDAVLYGEGCPWGRFGAALTVLGDVNGDKLTDVVGAPGEENRGAIVLP 556  
QY 541 HGTSGSGISPSHSORIASKSLSPLOYGQSLSGGQDITMDGLVDLTVGAQSHVLLRSQ 600  
DB 557 HGVLGPSISPSHSORIASQSSRLQYFGQALSGGQDLTDGLVDLAVGARGQVLLRTR 616  
QY 601 PVLRYKALMEFNPREVARNFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLWVGUSMQFIPAEIPSAFECREQVSVBQTLVQSNICLIDKRSKNLLGSRDLQSSVT 676

QY 661 YDLALDSCRPHSRVAVNFNETKSTRQTVGLTQTCETLKLQLPNCIEDPVPSPVILRLNF 720  
DB 677 LDALDPGRLSPRATFQETKNRSLSRVRVLGLKAHCENFNLLPSCVEDSVPTILRLNF 736  
QY 721 SLVGTPLSAFGNLRPLVAEDAQRULTALFPEKKNCGNDNICODDISITFSFMSLDCLVVG 780  
DB 737 TLVGKPLLAFLRNLRLPMLAALAQRYFTASLPPEKKNCGADHICQDNLGISFSFPLKSLVG 796  
QY 781 GPREFNVTVTVRNDGEDSYRTQVTFPFDLDSYKRVSTLQNRQSRQSRWLACESASSTEV 840  
DB 797 SNLEINAEVWVNDGEDSYGTTITFSPAGLSYRYVABGQKQGRSLRSLHLTCDSPVVG-- 854  
QY 841 SGALKSTSCSINHPIFENSEVETNITFDVDSKASLGNKLLKANVTSENMMPTNKTEF 900  
DB 855 SQGTWSTSCRINHILIFRGGAGITELATFDVSPKAVLGDRLLLTANVSSENNTPTSKTTF 914  
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTAS-ENTSRVMOHQYQVNSLQORSILPISLVFLVP 959  
DB 915 QLELPVKYAVYTVVSSHEQFTKYLNFSESEKESHVAMHRYQVNNLQORDLPVSNFVWP 974  
QY 960 VRLNQTVIMDRPQVTFSENLSTCHTKERLPSPSHDFLAELRKAPVAVNCSTAVCORIOCDI 1019  
DB 975 VELNQEAVMMDVEVSHPEQNESLRCSSEKIAPPASDFLAHIQKNPVLDCSIAGCLRPRCDV 1034  
QY 1020 PFTGQIEFENATLKNLSFDWYIKTSHNHLIIVSTABILEFENDSVFTLLPGQGFVRQTE 1079  
DB 1035 PSFVQBELDTLKNLSFGWVRQILQKVVSVVSAEITFTDTSVYSQLPQGEAFMRAQTT 1094  
QY 1080 TKVEPEFVNPRLPIYGVSSVGGLLALALITAAALYKLGFFKRYQKDMME 1128  
DB 1095 TVLEKYKVNFTPLIYGVSSIGGULLLALITAVLYKVGFFKRYQKEMEE 1143

## RESULT 12

US-09-350-259-4  
; Sequence 4, Application US/09350259  
; Patent No. US20020062008A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. US20020062008A1el Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/350,259  
; FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: 09/193,043  
; EARLIER FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03  
; NUMBER OF SEQ ID NOS: 114  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 1163  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-259-4

Query Match 58.7%; Score 3450; DB 9; Length 1163;  
Best Local Similarity 61.0%; Pred. No. 5.3e-282;  
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

QY 1 FNLDTENAMTFOENARFGOSVVOQLQGSRRVVGAPQBIIVANORGSILYQCDYSGSCEPI 60  
DB 20 FNLDTTELTAFRVDSAGFSDVSVQYANSWVVGAPQKITAANQTGGLYQCGYSTGACEPI 79  
QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLACGPTVHQTCSNTYVKGFLFSGNLRQOPQK 120  
DB 80 GLQVPEAVNMSLGLSLASTTSPPQLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTOR 137

```

; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1163
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-4

Query Match      58.7%; Score 3450; DB 10; Length 1163;
Best Local Similarity 61.0%; Pred. No. 5.3e-282;
Matches 689; Conservative 144; Mismatches 290; Indels 6; Gaps 4;

Qy 1 FNDTENAMTFQENARGFGOSVVLQGSVVVGAQPIVAANQKRSGLYQCDYSTGSCPEI 60
Db 20 FNDTEELTAFRVDSAGFGSVVQYANSVWVGAQPIVAANQKRSGLYQCDYSTGSCPEI 79
Qy 61 RLQVPEAVNMSGLSLAATTPPQLLACGPTVHQTCSNTYVGLCPFLPSNLRQDQK 120
Db 80 GLQVPEAVNMSGLSLAATTPPQLLACGPTVHQTCSNTYVGLCPFLPSNLRQDQK 137
Qy 121 FPEALRCQPEDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQLKSKTFLSMQYSEEF 180
Db 138 LPVSRQCEPRQEQDIVFLIDGSGSISRRNFATMNFVRAVISQFORPSTQSLQFSNKF 197
Qy 181 RIHTEFEFQNNPRSLVKPITQLGRTHPTATGVKRVIRELLNITNGARKNAFKILIVI 240
Db 198 QTHTEFEFRRTSNPLSLASVHQLQGTYYTATQNVVHRLFHASYGARRDAIKILIVI 257
Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDHVQVN 300
Db 258 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDHVQVN 317
Qy 301 NFEALKTQNLQREKIPAIETGOTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 318 DFDALKDQNLQREKIPAIETGOTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 377
Qy 361 GVFLYTSKESKSTFIMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFIMNSQENVMDRDSYLGYSTELALWKGVQSLVGLGAPRYQHIGLVAMFR 437
Qy 421 QNTGWNESNANVKGTOIGAYFGASCSVDVDSNGSTDLVIGAPHYYQTRGGQVSCPL 480
Db 438 QVSRQWRMKAIEVIGTQISGFYFASCSVDVDSNGSTDLVIGAPHYYQTRGGQVSCPL 497
Qy 481 PRGQARWQCDVAVLYGEOQPMGRFGAALTVDGVNGDKLTDVAIGAPCEENRGAVYLF 540
Db 498 PRGWR-RWMCDAVLYGEOQPMGRFGAALTVDGVNGDKLTDVAIGAPCEENRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVDITVGAQGHVLLRSQ 600
Db 557 HGVLGSPISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVDITVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFPNPREVARNVFCNDQVVGKEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLWVGVSQMFIPEIPRSAFEQVSEQLTVQSNICLYIDKRSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSAFVNETKSTRRTQTVGLGLTQTCETLKLQLPNCIEDPVSIVLRNF 720

```

```

Qy 121 FPEALRCQPEDSDIAFLIDGSGSIIIPDPRMKEFVSTVMEQLKSKTFLSMQYSEEF 180
Db 138 LPVSRQCEPRQEQDIVFLIDGSGSISRRNFATMNFVRAVISQFORPSTQSLQFSNKF 197
Qy 181 RIHTEFEFQNNPRSLVKPITQLGRTHPTATGVKRVIRELLNITNGARKNAFKILIVI 240
Db 198 QTHTEFEFRRTSNPLSLASVHQLQGTYYTATQNVVHRLFHASYGARRDAIKILIVI 257
Qy 241 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDHVQVN 300
Db 258 TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPRSEKSRQELNTIASKPRDHVQVN 317
Qy 301 NFEALKTQNLQREKIPAIETGOTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 360
Db 318 DFDALKDQNLQREKIPAIETGOTGSSSSFEHMSQEGFSAITNSGPLLSTVGSYDWAG 377
Qy 361 GVFLYTSKESKSTFIMTRVDSMDNDAYLGAAAIILNRVQSLVGLGAPRYQHIGLVAMFR 420
Db 378 GAFLYPPNMSPTFIMNSQENVMDRDSYLGYSTELALWKGVQSLVGLGAPRYQHIGLVAMFR 437
Qy 421 QNTGWNESNANVKGTOIGAYFGASCSVDVDSNGSTDLVIGAPHYYQTRGGQVSCPL 480
Db 438 QVSRQWRMKAIEVIGTQISGFYFASCSVDVDSNGSTDLVIGAPHYYQTRGGQVSCPL 497
Qy 481 PRGQARWQCDVAVLYGEOQPMGRFGAALTVDGVNGDKLTDVAIGAPCEENRGAVYLF 540
Db 498 PRGWR-RWMCDAVLYGEOQPMGRFGAALTVDGVNGDKLTDVAIGAPCEENRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVDITVGAQGHVLLRSQ 600
Db 557 HGVLGSPISPSHSQRIAGSKLSPRIQYFGQSLGGQDLTMDGLVDITVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFPNPREVARNVFCNDQVVGKEAGEVRVCLHVQKSTRDLRREGQIQSVVT 660
Db 617 PVLWVGVSQMFIPEIPRSAFEQVSEQLTVQSNICLYIDKRSKNLLGSRDLQSSVT 676
Qy 661 YDLALDSGRPHSAFVNETKSTRRTQTVGLGLTQTCETLKLQLPNCIEDPVSIVLRNF 720
Db 677 LDALAPGRSLPRAIFQETKRSLSRVKVLGKAKHCFNLLPSCVEDSVIPILRLNF 736
Qy 721 SLVCTPLSAFNLPRVLAEDAQRLTALFPPEKNCNDNIQDDLSITFSFMSDCLVVG 780
Db 737 TLVGKPLLAFLNRLPMLAALQRYFTASLPPEKNCADHIQDNLGIFSFPGLKSLIVG 796
Qy 781 GPRFNVTVVRNDEGSYRTQVTFPFLDLSYRKVSTLQNRQSORSWRLACESASTEV 840
Db 797 SNLELNAEVMVWNGEDSYGTTITFESHAPGLSYRYVABEQKQQLRSLHLTC--CSAFVG 854
Qy 841 SGALKSTSCSINHPIPPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF 900
Db 855 SQGTWSTSCRINHLIFRGAQITFLATEDVSPKAVGLDRLLLIANVSNNIPRTSKTIF 914
Qy 901 QLELPVKYAVMVVTSHGVTSTKYNLFTAS-ENTSRVMOHQYQVSNLQORSLPISLVFLVP 959
Db 915 QLELPVKYAVMVVTSHGVTSTKYNLFTAS-ENTSRVMOHQYQVSNLQORSLPISLVFLVP 974
Qy 960 VELNQTIVWDRPQVTFSENLSTCHTKERLPKSHSPFLAELRKAPVNCVSIAVCORIQCDI 1019
Db 975 VELNQEAVMVDEVSHPPQNPSPSRCSSEKTIAPASDFLAHQNPVLDSCSIAGLRCFCDV 1034
Qy 1020 PFGIOEBFNATLKNLSPDWIKTSHNHLITVSTAEILFNDVSFTLLPQGAFFRSQTE 1079
Db 1035 PSFSVQEBLDFTLKGNLSEFWRQILQKVSVVSVVAEIIIDTSTVSYQLPQGAFFRAQTI 1094
Qy 1080 TKVEPEVNPPLIIVGSSVGLLILALITAAALYKLGFFKQYKDMWSE 1128
Db 1095 TVLEKYKVNPIPLIVGSSIGGLLILALITAVLYKVGFFKQYKEMWEE 1143

RESULT 13
US-09-891-943-4
; Sequence 4, Application US/09891943

```



Db 677 LDALAPRLSPRAIFQETKRNLSRVRVLGKHCENFNLLPSCVEDSVIPIILRLNF 736  
Qy 721 SLVGTPLSAFAGNLRVLAEDAQRFTALFPPEKNCNDNI CODDLSITSPMSLDCLVVG 780  
Db 737 TLVGKPLAFNLRMLAALQRYETASLPPEKNCADHI CODNIGISFSPGLKSLVVG 796  
Qy 781 GPRFNVTVVRNDEGDSYRTQVTPFFFLDLSYRKVSTLQORSORSWELACESASSTEV 840  
Db 797 SNLEINAEVWMDGSDSYGTTIIFSHPAGLSYRVVAEGQKQGLRSLHLC--CSAPVG 854  
Qy 841 SGALKSTSCSINHPIFPENSEVFNITFDVDSKASLGKLLKANVTSENMPRTNKTEF 900  
Db 855 SQGTWSTSCRINHPIFRGAQITFLATPDVSPKAVGLDRLLIANVSSENNIPRTSKTIF 914  
Qy 901 QLELPVKYAVVMVYTHSGVSTKYLNFTAS-ENTSRVMOHQVOVNLGORSIPISLVFLVP 959  
Db 915 QLELPVKYAVVYVSSHEQFTYLNFSSEKESHVAMHRYQVNNLQORDLPVSEINWVP 974  
Qy 960 VRLNQTVIWRDPQVTFSENLSSSTCHTKERLPSPHSDFLAELRKAPVWNCIAVQCRIODI 1019  
Db 975 VELNQEAVWMDVEVSHQNPNSLRCSSEKIAPPASDFLAHQKNPVLDCSIAGCLURFRCDV 1034  
Qy 1020 PFFGIOEFNATLKGNSLSPDWIKTSHNHLIVSTAEILFNDVSFTLLPGQAFVRSQTE 1079  
Db 1035 PSFSVQBELDFTLKGNSLFGWVRQILQKYSVSVSAEIIIFDTSVYSQLPQGAEMRAQTI 1094  
Qy 1080 TKVEPFEVNPPLPIVGVSSVGGILLLALITAAALYKLGFFKROKDMME 1128  
Db 1095 TVLEKYKVNPIPLIVGVSSIGGLLLLALITAVLYKGVFFKROKEMEE 1143

## RESULT 14

US-09-350-259-2

; Sequence 2, Application US/09350259

; Patent No. US20020062008A1

; GENERAL INFORMATION:

; APPLICANT: Gallatin, Michael W.

; APPLICANT: Van der Vieren, Monica

; TITLE OF INVENTION: No. US20020062008A1el Human 2

; FILE REFERENCE: 27866/35004

; CURRENT APPLICATION NUMBER: US/09/350, 259

; EARLIER FILING DATE: 1999-07-08

; EARLIER APPLICATION NUMBER: 09/193,043

; EARLIER FILING DATE: 1998-11-16

; EARLIER APPLICATION NUMBER: 08/173,497

; EARLIER FILING DATE: 1993-12-23

; EARLIER APPLICATION NUMBER: 08/286,889

; EARLIER FILING DATE: 1994-08-05

; EARLIER APPLICATION NUMBER: 08/362,652

; EARLIER FILING DATE: 1994-12-21

; EARLIER APPLICATION NUMBER: 08/943,363

; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 1161

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-350-259-2

## Query Match

Best Local Similarity 58.2%; Score 3417; DB 9; Length 1161;

Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

Qy 1 FNLDTENAMTFQENARGGOSVQVQGGSRVVVGAPQEIIVAAHQSLYQCDYSTGSCPTI 60  
Db 17 FNLDEEPTIFQEDAGGQSVQVQGGSRVVVGAPLEVVAAHQSLYQCDYSTGSCPTI 76  
Qy 61 RLQVPVEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVVKGLCFPLFGNSRQPOK 120  
Db 77 PLHIRPEAVNMSLGLSLAATSPPOLLAGCPVTHQTCSENTYVVKGLCFPLFGNSRQPOK 135



Tue Nov 9 12:56:12 2004

```

; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; TITLE OF INVENTION: No. US2003007728A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1161
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-891-943-2

Query Match      58.2%; Score 3417; DB 10; Length 1161;
Best Local Similarity 59.8%; Pred. No. 3.3e-279;
Matches 675; Conservative 162; Mismatches 284; Indels 8; Gaps 6;

Qy 1 FNLDTENAMTPQENARFGQSVQIQGSRVVVGAPOEIVAAANORGSLYQCDYSTGSCPEI 60
Db 17 FNLDSVEPTIPQEDAGGFGQSVVQFGSRLVVGAPLEVVAAANTGRLYDCAATGMCQPI 76

Qy 61 RLQVPVEAVNMGLSLAATSPQLLAGCTVHTQTSENTYVVKGLCFGLFSNLRQOPQK 120
Db 77 PLHIRPEAVNMGLSLAATSTNGSRLLAGCTVHTQTSENTYVVKGLCFGLFSNLRQOPQK 135

Qy 121 FPEALRGCPQEDSDIARLIDGSGTIPHDPRRMKEFEVTVMEQKKSKTLFLSLMOYSEEF 180
Db 136 VPDATPECPQEMDVLFLIDGSGTDQNDNFQMGKFGVQAVMGQFEGTDTLFLALMOYSNLL 195

Qy 181 RIHFTFKFONNPRSLVKPITQLLGRHTHTATGVRKVRIRRELLNTNGARKNAKILIVI 240
Db 196 KIHTFTQRTSPSQSLVDPIVLQKGLTFTATGILTVTVIQLFHHKNGARKSAKKILIVI 255

Qy 241 TDGEKFGDPLGHEDVPIPADREGVIRYVIGDAPRSEKSOELNTIASKPRDHVFOVN 300
Db 256 TDGQYKDPLEYSYDIPQAEKAIIRYAGVGHAFQGTARQELNTISAPPQDHVFKVD 315

Qy 301 NFEALKTIONLREKIFALEGTQTGSSSFHEMSQEGFSAAITSNGPLLLSTVGSYDWAG 360
Db 316 NPAALGSIQKQLEKIYAVEGTQSRASSSFQHEMSQEGFSTALTMDGLFLGAVGFSWSG 375

Qy 361 GVLVYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRNVQSLVILGAPRYQHIGLVAMPR 420
Db 376 GAFLYPPNMPPTFINMSQENVMDRSDYLSYSTELALWKGVQNLVLGAPRYQHTGKAVIPT 435

Qy 421 QNTGWMESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLIGAPHYVEOTRGQSVSCPL 480
Db 436 QVSRQWRKKAETVGTQIGSYFGASLCSVDVDSNGSTDLLIGAPHYVEOTRGQSVSCPL 495

Qy 481 PRGORARWOCDAVLYGEOQOPWGRFGAALTVLGDVNGDKLTDVAIGAPGEEDNRGAVYLF 540
Db 496 PRGORVQWOCDAVLYGEOQOPWGRFGAALTVLGDVNGDKLTDVAIGAPGEENRGAVYLF 555

Qy 541 HGTSGSGISPSHSQRIAGSKLSRPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLLRQ 600
Db 556 HGASESGISPSHSQRIAGSKLSRPLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLLRQ 615

Qy 601 PVLRVKAINMEFNPREVARNVFECDQVVKGEAGEVRCILHYCKSTKDRRLREGQIOSVVT 660
Db 616 PVLKVGAMRFSPEVAVYRCWEKPSALEAGDATVCLTIQKSSLDQL--GDIQSSVR 673

Qy 661 YDLALDSGRPHSRVAFNETKNSRROTQVLGLTQTCETLKLQLPNCIEDPVSFVILRLNF 720

```

Search completed: November 9, 2004, 12:47:03  
Job time : 102.5 secs

```

674 FDLALDPGLTGRAIPNETKNPTLTRKTYLGLGHCETLKLPLPCDVEDVSPILLHLNF 733
Qy 721 SLVGTPLSAFGLNLRPVLAEDAQRLLFTALPFFKFNCGNDNICQDDLSTFMSLDCLVVG 780
Db 734 SLVREFIPSPQNLRFVLAVGSQDLFTASLPFFKFNCGQDGLCEGLGVTLSFSGLTITVG 793
Qy 781 GPREFNVTVVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNRQSRQSWRLACBSASSTEV 840
Db 794 SSLELNVITVMNAGSDSYGTVVSLYYPAGLSHRRVSGAQKQPHQSALRIACETV-PTD 852
Qy 841 SGALKSTCSINHPPIFENSEVTNITEDVDSKASLGNKLLKANKVNTSENMMPTNKTEF 900
Db 853 EG-LRSSRCVNHPIFHEGSGNGTFIVTFDVSYKATLGDRLMRASASSENKASSKATF 911
Qy 901 QLELPVKYAVYVMVVTSHGVSTKYLNFTASENTSRVMQHOYQVSNLQQRSLPISLVFLVP 959
Db 912 QLELPVKYAVYTMISRQESTKYFNFATSDKKMKEAHEHRYRVNLSQORDLAISINFWVP 971
Qy 960 VRLNQTVINDRPOQVTFSENLSSTCHTKERLPSSHDSDFLAELRKAPVNVCSIAVCORIOCDI 1019
Db 972 VLLNGVAVDVMVMEAPSQL--PCVSEKPPQSHSDFLTQISRSPMLDCSIADCLQFRCDV 1029
Qy 1020 PFGCIQBEFNATLKGNLSPDWYIKTSHNHLIIVSTAILEFNDVSVFTLLPQGGAFVRSQTE 1079
Db 1030 PSFSVOEELDTLLKGNLSPGWVRETLOKKVLVSVVAEITDTSVYSQLPQGEAFPMRAQME 1089
Qy 1080 TKVEPPFVNPPLPLIVGSSVGGLLLALITAAALYKLGFFKRYQKDMWSE 1128
Db 1090 MVLEDEEVYNAIPIIMGSSVGALLLALITATLYKLGFFKRYKEMLEMD 1138

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 9, 2004, 11:57:57 ; Search time 23 Seconds  
(without alignments)  
3278.416 Million cell updates/sec

Title: US-09-902-481B-6  
Perfect score: 5875  
Sequence: 1 FNLDTENAMTFQENARGFGQ.....FKROYKDMGEGPGGAEPPQ 1137

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5868	99.9	1153	1 US-08-173-497-3	Sequence 3, Appli
2	5868	99.9	1153	1 US-08-286-889-3	Sequence 3, Appli
3	5868	99.9	1153	1 US-08-485-618-3	Sequence 3, Appli
4	5868	99.9	1153	1 US-08-362-652-3	Sequence 3, Appli
5	5868	99.9	1153	2 US-08-605-672-3	Sequence 3, Appli
6	5868	99.9	1153	2 US-08-482-293A-3	Sequence 3, Appli
7	5868	99.9	1153	2 US-08-943-363-3	Sequence 3, Appli
8	5868	99.9	1153	3 US-09-193-043-3	Sequence 3, Appli
9	5868	99.9	1153	4 US-09-688-307A-3	Sequence 3, Appli
10	5868	99.9	1153	4 US-09-350-259-3	Sequence 3, Appli
11	5837.5	99.4	1152	2 US-08-476-062A-43	Sequence 3, Appli
12	5837.5	99.4	1152	5 PCT-US96-01314-43	Sequence 43, Appl
13	5837.5	99.4	1152	6 5424399-2	Sequence 43, Appl
14	3473	59.1	1163	2 US-08-476-062A-44	Patent No. 5424399
15	3473	59.1	1163	5 PCT-US96-01314-44	Sequence 44, Appl
16	3450	58.7	1163	1 US-08-173-497-4	Sequence 44, Appl
17	3450	58.7	1163	1 US-08-286-889-4	Sequence 4, Appli
18	3450	58.7	1163	1 US-08-485-618-4	Sequence 4, Appli
19	3450	58.7	1163	1 US-08-362-652-4	Sequence 4, Appli
20	3450	58.7	1163	2 US-08-605-672-4	Sequence 4, Appli
21	3450	58.7	1163	2 US-08-482-293A-4	Sequence 4, Appli
22	3450	58.7	1163	2 US-08-943-363-4	Sequence 4, Appli
23	3450	58.7	1163	3 US-09-193-043-4	Sequence 4, Appli
24	3450	58.7	1163	4 US-09-688-307A-4	Sequence 4, Appli
25	3450	58.7	1163	4 US-09-350-259-4	Sequence 4, Appli
26	3417	58.2	1161	1 US-08-173-497-2	Sequence 2, Appli
27	3417	58.2	1161	1 US-08-286-889-2	Sequence 2, Appli

28	3417	58.2	1161	1 US-08-485-618-2	Sequence 2, Appli
29	3417	58.2	1161	1 US-08-362-652-2	Sequence 2, Appli
30	3417	58.2	1161	2 US-08-605-672-2	Sequence 2, Appli
31	3417	58.2	1161	2 US-08-482-293A-2	Sequence 2, Appli
32	3417	58.2	1161	2 US-08-943-363-2	Sequence 2, Appli
33	3417	58.2	1161	3 US-09-193-043-2	Sequence 2, Appli
34	3417	58.2	1161	4 US-09-688-307A-2	Sequence 2, Appli
35	3417	58.2	1161	4 US-09-350-259-2	Sequence 2, Appli
36	3401.5	57.9	1161	1 US-08-485-618-99	Sequence 99, Appl
37	3401.5	57.9	1161	2 US-08-605-672-99	Sequence 99, Appl
38	3401.5	57.9	1161	2 US-08-482-293A-99	Sequence 99, Appl
39	3401.5	57.9	1161	2 US-08-943-363-99	Sequence 99, Appl
40	3401.5	57.9	1161	3 US-09-193-043-99	Sequence 99, Appl
41	3401.5	57.9	1161	4 US-09-688-307A-99	Sequence 99, Appl
42	3401.5	57.9	1161	4 US-09-350-259-99	Sequence 99, Appl
43	3239.5	55.1	1161	3 US-09-193-043-55	Sequence 55, Appl
44	3239.5	55.1	1161	4 US-09-688-307A-55	Sequence 55, Appl
45	3239.5	55.1	1161	4 US-09-350-259-55	Sequence 55, Appl

## ALIGNMENTS

RESULT 1  
US-08-173-497-3  
; Sequence 3, Application US/08173497  
; Patent No. 5437958  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van Der Vieren, Monica  
; TITLE OF INVENTION: No. 5437958el Human 2 Integrin Alpha  
; TITLE OF INVENTION: Subunit  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 S. Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/173,497  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5437958and, Greta E.  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/31363  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-173-497-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 FNLDTENAMTFQENARGFGQSVVQLQSRVVVGAPQBIIVANQSGSLYQCDYSTGSCEPI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLQSRVVVGAPQBIIVANQSGSLYQCDYSTGSCEPI 76

RESULT 2  
US-08-286-889-3  
; Sequence 3, Application US/08286889  
; Patent No. 5470953  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Mich  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 51  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/286,889  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: P38,659  
; REFERENCE/DOCKET NUMBER: 27865/32168  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-286-889-3  
Query Match 99.9%; Score 5868; DB 1; Length 1153;  
Best Local Similarity 99.6%; Pred No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 61 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTVYVKGCLFLFGSNLRQPOK 120  
Db 77 RLQVPVEAVNMSLGLSLAATSPQLLACGPTVHQTCSNTVYVKGCLFLFGSNLRQPOK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEOLKSKTLFSLMQYSEEF 180  
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEOLKSKTLFSLMQYSEEF 196  
QY 181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELNTINGARKNAFKILVI 240  
Db 197 RHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVVRELENTINGARKNAFKILVI 256  
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN 300  
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN 316  
QY 301 NFPEALKTIONLREKIFAEGTGTGSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 360  
Db 317 NFPEALKTIONLREKIFAEGTGTGSSSFEHMSQEGFSAATISNGPLLSITVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAIILNRRVQSLVGLGAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGQVSYCPL 480  
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVIGAPHYYEOTRGQVSYCPL 496  
QY 481 PRGORARWQCDVILYGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVLYF 540  
Db 497 PRGORARWQCDVILYGQGPWGRFGAALTVLGVNGDKLTDVAIGAPGEDNREGAVLYF 556  
QY 541 HGTSGSGISPSHOSRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHOSRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
QY 601 PVLRVKAIMFPNPREVARNFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 617 PVLRVKAIMFPNPREVARNFECDNQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETKSTRQTOVLGTTQTCETLKLQLPNCIEDPVSPIVRLNF 720  
Db 677 YDLALDSGRPHSRVAFNETKSTRQTOVLGTTQTCETLKLQLPNCIEDPVSPIVRLNF 736  
QY 721 SLVGTPLSAFGNLRVLAEDAQRLEFALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
Db 737 SLVGTPLSAFGNLRVLAEDAQRLEFALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
QY 781 GPREFNVTVRNDGEDSVRTQVTFPFDLISYRKVSTLONORSORSWRLACESASSTEV 840  
Db 797 GPREFNVTVRNDGEDSVRTQVTFPFDLISYRKVSTLONORSORSWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPNRTKTEF 900  
Db 857 SGALKSTSCSINHPIFPENSEVTENITFDVDSKASLGNKLLKANVTSENMPNRTKTEF 916  
QY 901 QLELPVKAVYVMTVSHGVSTKYLNFPTASENTSRVMQHOYQVSNLQORSILPISLPLVPV 960  
Db 917 QLELPVKAVYVMTVSHGVSTKYLNFPTASENTSRVMQHOYQVSNLQORSILPISLPLVPV 976  
QY 961 RLNQTVIMDRPOVTPSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1020  
Db 977 RLNQTVIMDRPOVTPSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCORIQCDIP 1036  
QY 1021 FFGIQEEFNATLKNLSFDWYIKTSNNHLLIVSTAILFNDSVFTLLPGQAFVRSQTEI 1080  
Db 1037 FFGIQEEFNATLKNLSFDWYIKTSNNHLLIVSTAILFNDSVFTLLPGQAFVRSQTEI 1096  
QY 1081 KVEPPEFVNPPLPIVGSVGLLLALITAAALYKLGFFFKQYKDMMSGGPPGABPQ 1137  
Db 1097 KVEPPEFVNPPLPIVGSVGLLLALITAAALYKLGFFFKQYKDMMSGGPPGABPQ 1153

QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLAGPYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLAGPYQHIGLVAMFR 436  
QY 421 QNTGWESNANVKGTOIGAYGASLCSVDVDSNGSTDLVLGAPHYHQTREGGVSVCP 480  
Db 437 QNTGWESNANVKGTOIGAYGASLCSVDVDSNGSTDLVLGAPHYHQTREGGVSVCP 496  
QY 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWQCDVLYGEGQDPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDTLVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDTLVGAQGHVLLRSQ 616  
QY 601 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
Db 617 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQPCNIEDPSPVILRLNF 720  
Db 677 YDLALDSGRPHSRVAFNETKNSTRQTVGLTQTCETLKLQPCNIEDPSPVILRLNF 736  
QY 721 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSITFMSLDCLVVG 780  
Db 737 SLVGTPLSAFGLNRPVLAEDAQRLLFTALFPFEKNCNDNICODDLSITFMSLDCLVVG 796  
QY 781 GPREFNVTVVRNDGDSYQTVTFPPPLDLSYRKVSTLQNRQSQRSLACASSTEV 840  
Db 797 GPREFNVTVVRNDGDSYQTVTFPPPLDLSYRKVSTLQNRQSQRSLACASSTEV 856  
QY 841 SGALKSTSCSINPIPIPENSEVFNFTFDVDSKASLGNKLLKXANTVSENMPRTNTEF 900  
Db 857 SGALKSTSCSINPIPIPENSEVFNFTFDVDSKASLGNKLLKXANTVSENMPRTNTEF 916  
QY 901 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLPV 960  
Db 917 QLELPVKYAVVMVTSHGVSSTKYLNFTASENTSRVMOHQYQVSNLQORSIPISLVFLPV 976  
QY 961 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1020  
Db 977 RLNQTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCORIQCIP 1036  
QY 1021 FFGIOEEFNATLKNLSFDWYIKTSHNLLIYSTABILFNDVSFTLLPGQAFVRSQTET 1080  
Db 1037 FFGIOEEFNATLKNLSFDWYIKTSHNLLIYSTABILFNDVSFTLLPGQAFVRSQTET 1096  
QY 1081 KVEPEFVNPPLPLVGVSSVGLLLALITAAALYKLGFFKRYQKDMMEGGPPGAEPO 1137  
Db 1097 KVEPEFVNPPLPLVGVSSVGLLLALITAAALYKLGFFKRYQKDMMEGGPPGAEPO 1153

## RESULT 3

US-08-485-618-3  
; Sequence 3, Application US/08485618  
; Patent No. 5728533  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; City: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/485,618  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32797  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-485-618-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAAQSGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQCSTYVYKGLCFIFGSLRQOPQK 120  
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQCSTYVYKGLCFIFGSLRQOPQK 136  
QY 121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEOLKSKTKLFSLMQYSEEF 180  
Db 137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEOLKSKTKLFSLMQYSEEF 196  
QY 181 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIIRLLNITNGARKNAFKILIVI 240  
Db 197 RHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIIRLLNITNGARKNAFKILIVI 256  
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQEIANTASKPRDHVQVN 300  
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDGAFRSEKSRQEIANTASKPRDHVQVN 316  
QY 301 NFEALKTIONQUREKIFAIEGTQTGSSSFHEHMQEGFSAAITNSGPLLSTVGSVDWAG 360  
Db 317 NFEALKTIONQUREKIFAIEGTQTGSSSFHEHMQEGFSAAITNSGPLLSTVGSVDWAG 376  
QY 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLAGPYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVQSLVLAGPYQHIGLVAMFR 436  
QY 421 QNTGWESNANVKGTOIGAYGASLCSVDVDSNGSTDLVLGAPHYHQTREGGVSVCP 480  
Db 437 QNTGWESNANVKGTOIGAYGASLCSVDVDSNGSTDLVLGAPHYHQTREGGVSVCP 496  
QY 481 PRGQARWQCDVLYGEGQDPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWQCDVLYGEGQDPWGRFGAALTVDGVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDTLVGAQGHVLLRSQ 600

REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32391  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-362-652-3

Query Match 99.9%; Score 5868; DB 1; Length 1153;  
Best Local Similarity 99.8%; Pred. No 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTFQENARGFGQSVVLOGSRVVVGAPQEIIVAAQORGSIIYQCDYSTGSCPEI 60  
Db 17 FNLDTENAMTFQENARGFGQSVVLOGSRVVVGAPQEIIVAAQORGSIIYQCDYSTGSCPEI 76  
QY 61 RLQVPVAVNMISLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLSGNLRQOQOK 120  
Db 77 RLQVPVAVNMISLGLSLAATTSPOLLACGPTVHQTCSNTYVKGCLFGLSGNLRQOQOK 136  
QY 121 FPEALRCPOEDSDIAFLIDSGSIIIPHDFRRMKFEVSTVMEQIKKSKTLFSLMOYSEEF 180  
Db 137 FPEALRCPOEDSDIAFLIDSGSIIIPHDFRRMKFEVSTVMEQIKKSKTLFSLMOYSEEF 196  
QY 181 RIHFTKFEFQNNPNRSLVKPIITQLLGRTHATGVRKIVRELNITNGARKNAFKILIVI 240  
Db 197 RIHFTKFEFQNNPNRSLVKPIITQLLGRTHATGVRKIVRELNITNGARKNAFKILIVI 256  
QY 241 TDGKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSHOELNTIASKPPRHHVFOVN 300  
Db 257 TDGKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSHOELNTIASKPPRHHVFOVN 316  
QY 301 NFEALKTIONOLREKIFAIEGTQTGSSSPHEMSOEGFSAAITNSGPLLSTVGSYDWAG 360  
Db 317 NFEALKTIONOLREKIFAIEGTQTGSSSPHEMSOEGFSAAITNSGPLLSTVGSYDWAG 376  
QY 361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRNVQSLVGLGAPRYQHIGLVAMFR 420  
Db 377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGAAAAIILNRNVQSLVGLGAPRYQHIGLVAMFR 436  
QY 421 QNTGMWESNANVKGTQICAYFGASLCSDVDNSGSTDLVILICAPHYYEOTRGQVSVCP 480  
Db 437 QNTGMWESNANVKGTQICAYFGASLCSDVDNSGSTDLVILICAPHYYEOTRGQVSVCP 496  
QY 481 PRGQARWQCDVAVLYGEOGPWGRFGAALTIVLGDVNGDKLTDVAITGAPGEEDNRGAVYLF 540  
Db 497 PRGQARWQCDVAVLYGEOGPWGRFGAALTIVLGDVNGDKLTDVAITGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTVAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTVAQGHVLLRSQ 616  
QY 601 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 660  
Db 617 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 676  
QY 661 YDLALDSGRPHSAVRNFETKSTRQTQVILGTQTCETLKLQPLNCIEDPVPSPVLRNLF 720  
Db 677 YDLALDSGRPHSAVRNFETKSTRQTQVILGTQTCETLKLQPLNCIEDPVPSPVLRNLF 736  
QY 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFPEKNCNDNICODDLSITFSPMSLCLVVG 780  
Db 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFPEKNCNDNICODDLSITFSPMSLCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSVRKVSTLQNRORSRWSLACESASSTEV 840

557 HGTSGSGISPSHSQRIAGSKLSPRLOYFGQSLSGQDLTMDGLVDLTVAQGHVLLRSQ 616  
QY 601 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 660  
Db 617 PVLRVKAIMENPREVARNVFECNDQVVKGEAGEVRVCLHVOKSTRDLREGOIQSVVT 676  
QY 661 YDLALDSGRPHSAVRNFETKSTRQTQVILGTQTCETLKLQPLNCIEDPVPSPVLRNLF 720  
Db 677 YDLALDSGRPHSAVRNFETKSTRQTQVILGTQTCETLKLQPLNCIEDPVPSPVLRNLF 736  
QY 721 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFPEKNCNDNICODDLSITFSPMSLCLVVG 780  
Db 737 SLVGTPLSAFAGNLRPVLAEDAQRLLFTALFPFPEKNCNDNICODDLSITFSPMSLCLVVG 796  
QY 781 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSVRKVSTLQNRORSRWSLACESASSTEV 840  
Db 797 GPREFNVTVVRNDGEDSYRTQVTFPPFLDLSVRKVSTLQNRORSRWSLACESASSTEV 856  
QY 841 SGALKSTSCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKXANTSENMMPTKNTKEF 900  
Db 857 SGALKSTSCSINHPFIPENSEVTFNITFDVDSKASLGNKLLKXANTSENMMPTKNTKEF 916  
QY 901 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMOHQVSNLQORSLSPLSLVFLVPV 960  
Db 917 QLELPVKYAVYVMVTSHGVS TKYLNFTASENTSRVMOHQVSNLQORSLSPLSLVFLVPV 976  
QY 961 RLNGTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1020  
Db 977 RLNGTVIWRDPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVNCIAVCQRIQCDIP 1036  
QY 1021 PFGIOEFNATLKNLSPDWTKVSHHLLIVSTAEILFNDVSFTLPGQGAFAVRSOTET 1080  
Db 1037 PFGIOEFNATLKNLSPDWTKVSHHLLIVSTAEILFNDVSFTLPGQGAFAVRSOTET 1096  
QY 1081 KVEPFEVNPPLVIGSSVGLLALITLTAALYKLGPFKROYKDMSEGGPPGAEPO 1137  
Db 1097 KVEPFEVNPPLVIGSSVGLLALITLTAALYKLGPFKROYKDMSEGGPPGAEPO 1153

RESULT 4  
US-08-362-652-3  
Sequence 3, Application US/08362652  
Patent No. 5766850  
GENERAL INFORMATION:  
APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No 5766850el Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 93  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/362,652  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,899  
FILING DATE: 5-AUG-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.

Db 797 GPREFNVTYVRNDEGDSYRTQVTFPPFLDLRYKSTVLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVVMQHQYQVSNLQGRSLPISLFLVFPV 960  
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVVMQHQYQVSNLQGRSLPISLFLVFPV 976  
QY 961 RLNQTVIWDROPVTFSENLSSCHTKERLPSPSDFLAELRKAPVNVCSIAVCQIQCDIP 1020  
Db 977 RLNQTVIWDROPVTFSENLSSCHTKERLPSPSDFLAELRKAPVNVCSIAVCQIQCDIP 1036  
QY 1021 PFGIQEENATLKGNSLFDWIKTSHNHLIIIVSTAELPNSDVFLLPQOGAFVRSOTET 1080  
Db 1037 PFGIQEENATLKGNSLFDWIKTSHNHLIIIVSTAELPNSDVFLLPQOGAFVRSOTET 1096  
QY 1081 KVEPFEVNPPLIVGSSVGGILLIALLITAALYKLGFFKQYKDMMSGGPPGAEPPQ 1137  
Db 1097 KVEPFEVNPPLIVGSSVGGILLIALLITAALYKLGFFKQYKDMMSGGPPGAEPPQ 1153

## RESULT 5

US-08-605-672-3  
; Sequence 3, Application US/08605672  
; Patent No. 581/515  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 581/515e1 Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/605,672  
; FILING DATE:  
; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856

; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; TYPE: amino acid  
; LENGTH: 1153 amino acids  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-605-672-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 ENLDTENAMTQENARGGQSVVQOGSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGCEPI 60  
Db 17 ENLDTENAMTQENARGGQSVVQOGSRVVVVGAPQEIIVAAQNRGSLYQCDYSTGCEPI 76  
QY 61 RLQVPEAVNMSLGLSLAATSPPOLLAAGPTVHOTCSNTYVVKGLCFGLFGLNLRQPOK 120  
Db 77 RLQVPEAVNMSLGLSLAATSPPOLLAAGPTVHOTCSNTYVVKGLCFGLFGLNLRQPOK 136  
QY 121 FPEALRGCPQEDSIAFLIDGSGSIIIPHDPRMKEFVSTVMEQLKSKTLESLMOYSEEF 180  
Db 137 FPEALRGCPQEDSIAFLIDGSGSIIIPHDPRMKEFVSTVMEQLKSKTLESLMOYSEEF 196  
QY 181 RIHETFEKFEFQNNPRLSLVKPITQLLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 240  
Db 197 RIHETFEKFEFQNNPRLSLVKPITQLLGRTHATGVKRVIRELLNITNGARKNAFKILIVI 256  
QY 241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAPRSEKSKOELENTIASKPPRDHVFQV 300  
Db 257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVGDAPRSEKSKOELENTIASKPPRDHVFQV 316  
QY 301 NFEALKTIQNLREKIFAIEGTQTGSSSPSEHEMSQEGFSAATITNSGSLTSTGSDYDAG 360  
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSPSEHEMSQEGFSAATITNSGSLTSTGSDYDAG 376  
QY 361 GVFLYTSKESKSTFNMTRVDSMDNDAYLGVAAILLNRVQSLVGLAPRYQHIGLVAMER 420  
Db 377 GVFLYTSKESKSTFNMTRVDSMDNDAYLGVAAILLNRVQSLVGLAPRYQHIGLVAMER 436  
QY 421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQSVVCP 480  
Db 437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSDTLVLIGAPHYEQTRGGQSVVCP 496  
QY 481 PRQARWQCDAVLYGEOGQPGWGRFGAALTGLVDVNGDKLTDVAIGAPGEEDNRGAVYLF 540  
Db 497 PRQARWQCDAVLYGEOGQPGWGRFGAALTGLVDVNGDKLTDVAIGAPGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ 600  
Db 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDTLTVGAQGHVLLRSQ 616  
QY 601 PVLRVKAI MEENPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQISVVT 660  
Db 617 PVLRVKAI MEENPREVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQISVVT 676  
QY 661 YDLALDSGRPHSRAVFNETKSTRQTQVLTQTCETLKLQLPNCIEDPVSPVLRINF 720  
Db 677 YDLALDSGRPHSRAVFNETKSTRQTQVLTQTCETLKLQLPNCIEDPVSPVLRINF 736  
QY 721 SLVGTPLSAFNGNLRPVLAEDAQRFTALFPPEKNCNDNIQDDLSITFSFMSLCLVVG 780  
Db 737 SLVGTPLSAFNGNLRPVLAEDAQRFTALFPPEKNCNDNIQDDLSITFSFMSLCLVVG 796  
QY 781 GPREENVTYVRNDEGDSYRTQVTFPPFLDLRYKSTVLQNRQSRWRLACESASSTEV 840  
Db 797 GPREENVTYVRNDEGDSYRTQVTFPPFLDLRYKSTVLQNRQSRWRLACESASSTEV 856  
QY 841 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 900  
Db 857 SGALKSTSCSINHPIPFENSEVTFNITFDVDSKASLGKLLKLLKANVTSENMPRTNKTEF 916  
QY 901 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVVMQHQYQVSNLQGRSLPISLFLVFPV 960  
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASENTRVVMQHQYQVSNLQGRSLPISLFLVFPV 976  
QY 961 RLNQTVIWDROPVTFSENLSSCHTKERLPSPSDFLAELRKAPVNVCSIAVCQIQCDIP 1020  
Db 977 RLNQTVIWDROPVTFSENLSSCHTKERLPSPSDFLAELRKAPVNVCSIAVCQIQCDIP 1036

1021 FFGIOEFNATLKGNSLSPWYIKTSHNHLLIYSTABILFNDVSFTLLPGGAFVRSQTET 1080  
1037 FFGIOEFNATLKGNSLSPWYIKTSHNHLLIYSTABILFNDVSFTLLPGGAFVRSQTET 1096  
1081 KVEPFVFPNPLPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1137  
1097 KVEPFVFPNPLPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1153

RESULT 6  
US-08-482-293A-3  
; Sequence 3, Application US/08482293A  
; Patent No. 5831029  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, W. Michael  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit  
; NUMBER OF SEQUENCES: 103  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sear Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/482,293A  
; FILING DATE:  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/173,497  
; FILING DATE: 23-DEC-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/286,889  
; FILING DATE: 5-AUG-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/362,652  
; FILING DATE: 21-DEC-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/32684  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1153 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-482-293A-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGCEPI 60  
17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVGAPQEIIVAANQSGSLYQCDYSTGCEPI 76  
61 RLQVPVEAVNMSGLSLAATTPSPQLACGPTVHQTCSNTYVYKGLCFLFGSNLRQOPQK 120  
77 RLQVPVEAVNMSGLSLAATTPSPQLACGPTVHQTCSNTYVYKGLCFLFGSNLRQOPQK 136  
121 FPEALRGCPQEDSIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLFSLMOYSEEF 180

137 FPEALRGCPQEDSIAFLIDGSGSIIPHDFRMKEFVSTVMEQLKSKTFLFSLMOYSEEF 196  
181 RHFTFKFQNNPNRSLVKPITQLLGRHTATGVRKVIRELINITNGARKNAFKLILVI 240  
197 RHFTFKFQNNPNRSLVKPITQLLGRHTATGIRKVVRELFNITNGARKNAFKLILVI 256  
241 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNITIASPPRDHVFQVN 300  
257 TDGEKFGDPLGYEDVIPEADREGVIRYVIGVDAPFRSEKSRQELNITIASPPRDHVFQVN 316  
301 NFEALKTIONQLREKIPALEGTQTGSSSSPEHEMSQEGSAAITSGNPILSTVGSYDWDAG 360  
317 NFEALKTIONQLREKIPALEGTQTGSSSSPEHEMSQEGSAAITSGNPILSTVGSYDWDAG 376  
361 GVFLYTSKEKSTFINMTRVDSDMNDAYLVYAAAIILNRNVQSLVGLGAPRYOHIGLVAMFR 420  
377 GVFLYTSKEKSTFINMTRVDSDMNDAYLVYAAAIILNRNVQSLVGLGAPRYOHIGLVAMFR 436  
421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTREGGOVSCPL 480  
437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYEOTREGGOVSCPL 496  
481 PRGORARWQCDVILYGEQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 540  
497 PRGORARWQCDVILYGEQGPWGRFGAALTVDGVNGDKLTDVAIGAPGEDNRGAVYLF 556  
541 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 600  
557 HGTSGSGISPSHSORIASGLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ 616  
601 PVLAVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
617 PVLAVKAIMEFNPREVARNVPECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
661 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 720  
677 YDLALDSGRPHSRVAFNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVSPIVRLNF 736  
721 SLVGTPLSAFGNLRPVLAEADAQRULTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 780  
737 SLVGTPLSAFGNLRPVLAEADAQRULTALFPPEKNCNDNIQDDLSITFSFMSLDCLVVG 796  
781 GPREFNVTVVRNDGEDSYRQTVTFPPDLDSYRKVSTLQNRORSORSWRLACESASSTEV 840  
797 GPREFNVTVVRNDGEDSYRQTVTFPPDLDSYRKVSTLQNRORSORSWRLACESASSTEV 856  
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRINKTEF 900  
857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGKLLKANVTSENMPRINKTEF 916  
901 QLELPVKYAVYVMVTSYHGVSTKYLNFTASENTSRVMOHQYQVSNLQGRSLPISLPLVVPV 960  
917 QLELPVKYAVYVMVTSYHGVSTKYLNFTASENTSRVMOHQYQVSNLQGRSLPISLPLVVPV 976  
961 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSAVCORIQCDIP 1020  
977 RLNQTVIWDPRQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNVCSAVCORIQCDIP 1036  
1021 FFGIOEFNATLKGNSLSPWYIKTSHNHLLIYSTABILFNDVSFTLLPGGAFVRSQTET 1080  
1037 FFGIOEFNATLKGNSLSPWYIKTSHNHLLIYSTABILFNDVSFTLLPGGAFVRSQTET 1096  
1081 KVEPFVFPNPLPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1137  
1097 KVEPFVFPNPLPLIVGSSVGGILLLALITAAALYKLGFFKQYKDMSEGPPGAEPO 1153

RESULT 7  
US-08-943-363-3  
; Sequence 3, Application US/08943363  
; Patent No. 5837478  
; GENERAL INFORMATION:



APPLICANT: Gallatin, W. Michael  
APPLICANT: Van der Vieren, Monica  
TITLE OF INVENTION: No. 5837478e1 Human 2 Integrin Alpha Subunit  
NUMBER OF SEQUENCES: 114  
CORRESPONDENCE ADDRESS:  
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun  
STREET: 233 South Wacker Drive, 6300 Sear Tower  
CITY: Chicago  
STATE: Illinois  
COUNTRY: United States  
ZIP: 60606-6402  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/943,363  
FILING DATE:  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/173,497  
FILING DATE: 23-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/286,889  
FILING DATE: 5-AUG-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/362,652  
FILING DATE: 21-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams Jr., Joseph A.  
REGISTRATION NUMBER: 38,659  
REFERENCE/DOCKET NUMBER: 27866/32684  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312-474-6300  
TELEFAX: 312-474-0448  
TELEX: 25-3856  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1153 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-943-363-3

Query Match 99.9%; Score 5868; DB 2; Length 1153;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 FNLDTENAMTQBNARFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 60  
DB 17 FNLDTENAMTQBNARFGQSVVQLQGSRRVVVGAPQEIIVAAORGSLYQCDYSTGSCPEI 76  
QY 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVKGICFLFGSNLRQOPQK 120  
DB 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHTQCSNTYVKGICFLFGSNLRQOPQK 136  
QY 121 FPEALRGCPQEDSDIAFLDGGSGIIPHPFRMKFEVSTVMQKLGKSTLFSIMQYSEEF 180  
DB 137 FPEALRGCPQEDSDIAFLDGGSGIIPHPFRMKFEVSTVMQKLGKSTLFSIMQYSEEF 196  
QY 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILVI 240  
DB 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILVI 256  
QY 241 TDGEKFGDPLGVEDVIPEADRGVIRYIVGVGDAFRSEKSRQELNTIASKPPRDHFQVN 300  
DB 257 TDGEKFGDPLGVEDVIPEADRGVIRYIVGVGDAFRSEKSRQELNTIASKPPRDHFQVN 316  
QY 301 NPEALKTIONQRKXIFATEGTQSSSFEHMSQEGFSAITNGPILLSTVGSDYDAG 360  
DB 317 NPEALKTIONQRKXIFATEGTQSSSFEHMSQEGFSAITNGPILLSTVGSDYDAG 376

QY 361 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIILNRVQSLVLCAPRYQHIGLVAMPR 420  
DB 377 GVFLYTSKEKSTFINMTVRVDSMDNDAYLGYAAAIILNRVQSLVLCAPRYQHIGLVAMPR 436  
QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYQTRGGQSVCP 480  
DB 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDLVLIGAPHYYQTRGGQSVCP 496  
QY 481 PRGORARWQCDVLYGEGQCPWGRFGAALTVLGVDVNGDKLTDVAIGAGEEDNRGAVYLF 540  
DB 497 PRGORARWQCDVLYGEGQCPWGRFGAALTVLGVDVNGDKLTDVAIGAGEEDNRGAVYLF 556  
QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAAQGHVLLRSQ 600  
DB 557 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGQDLTMDGLVDLTVGAAQGHVLLRSQ 616  
QY 601 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 660  
DB 617 PVLRVKAIMFENPREVARNVFECNDQVVKGEAGEVRVCLHVQKSTRDLREGQIQSVVT 676  
QY 661 YDLALDSGRPHSRAVFNETKNSTRROTQVLGLTQTCETLKLQPNCLIEDPVPVILRLNF 720  
DB 677 YDLALDSGRPHSRAVFNETKNSTRROTQVLGLTQTCETLKLQPNCLIEDPVPVILRLNF 736  
QY 721 SLVGTPLSAFNGRPLVLAEDAQRFTALFPPEKNCNDNICQDDLSITFFSMDLCLVVG 780  
DB 737 SLVGTPLSAFNGRPLVLAEDAQRFTALFPPEKNCNDNICQDDLSITFFSMDLCLVVG 796  
QY 781 GPREFNVTVVRNDGDSYETQVTFEPDLISYRKVSTLQNRQSRQSWRLACESASSTEV 840  
DB 797 GPREFNVTVVRNDGDSYETQVTFEPDLISYRKVSTLQNRQSRQSWRLACESASSTEV 856  
QY 841 SGALKSTCSINHIPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 900  
DB 857 SGALKSTCSINHIPIFPENSEVFNITFDVDSKASLGNKLLKANVTSENNMPTNKTEF 916  
QY 901 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSILPISLVLPV 960  
DB 917 QLELPVKYAVYVMTVSHGVSTKYLNFTASENTSRVMQHOYQVSNLQORSILPISLVLPV 976  
QY 961 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFTAEIRKAPVVCNCSIAVCORI 1020  
DB 977 RLNQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFTAEIRKAPVVCNCSIAVCORI 1036  
QY 1021 FFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTABILFNDVSFTLLPGGAFVRSQTET 1080  
DB 1037 FFGIQEEFNATLKGNSLFDWYIKTSHNLLIIVSTABILFNDVSFTLLPGGAFVRSQTET 1096  
QY 1081 KVEPEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1137  
DB 1097 KVEPEFVNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGGPPGAEPO 1153

RESULT 8  
US-09-193-043-3  
; Sequence 3, Application US/09193043  
; Patent No. 6251395  
; GENERAL INFORMATION:  
; APPLICANT: Gallatin, Michael W.  
; APPLICANT: Van der Vieren, Monica  
; TITLE OF INVENTION: No. 6251395e1 Human 2  
; FILE REFERENCE: 27866/35004  
; CURRENT APPLICATION NUMBER: US/09/193,043  
; CURRENT FILING DATE: 1998-11-16  
; EARLIER APPLICATION NUMBER: 08/173,497  
; EARLIER FILING DATE: 1993-12-23  
; EARLIER APPLICATION NUMBER: 08/286,889  
; EARLIER FILING DATE: 1994-08-05  
; EARLIER APPLICATION NUMBER: 08/362,652  
; EARLIER FILING DATE: 1994-12-21  
; EARLIER APPLICATION NUMBER: 08/943,363  
; EARLIER FILING DATE: 1997-10-03

; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: Patent In Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-193-043-3									
Query Match 99.9%; Score 5868; DB 3; Length 1153;									
Best Local Similarity 99.6%; Pred. No. 0;									
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI	60						
Db	17	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI	76						
Qy	61	RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLFGSNLRQOPQK	120						
Db	77	RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLFGSNLRQOPQK	136						
Qy	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSEEF	180						
Db	137	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSEEF	196						
Qy	181	RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIRELLNITNGARKNAFKILIVI	240						
Db	197	RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIRELLNITNGARKNAFKILIVI	256						
Qy	241	TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN	300						
Db	257	TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN	316						
Qy	301	NFEALKTIONOLREKIFAIEGTQTGSSSPHEHMSQEGFSAATISNGPLLSITVGSYDWAG	360						
Db	317	NFEALKTIONOLREKIFAIEGTQTGSSSPHEHMSQEGFSAATISNGPLLSITVGSYDWAG	376						
Qy	361	GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAILLRNVOSLVLGAPYOHIGLVAMFR	420						
Db	377	GVFLYTSKEKSTFINMTVDSMDNDAYLGYAAAILLRNVOSLVLGAPYOHIGLVAMFR	436						
Qy	421	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSCVPL	480						
Db	437	QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDVLVIGAPHYEQTRGGQVSCVPL	496						
Qy	481	PRGQARWQCDVAVLYGEOQPMGRFGAALTVDGVNGDKLTDVAICAGEEDNRGAVLYF	540						
Db	497	PRGQARWQCDVAVLYGEOQPMGRFGAALTVDGVNGDKLTDVAICAGEEDNRGAVLYF	556						
Qy	541	HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVBLTVGAQGHVLLRSQ	600						
Db	557	HGTSGSGISPHSQRISAGSKLSPRLQYFGQSLSGQDLTMDGLVBLTVGAQGHVLLRSQ	616						
Qy	601	PVLVRKALMEFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT	660						
Db	617	PVLVRKALMEFNPVARNVFCNDQVVKGEAGEVRVCLHVQKSTRDLREGQIOSVVT	676						
Qy	661	YDALDSGPHGRVAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPIVLRNF	720						
Db	677	YDALDSGPHGRVAVNETKNSTRQTVLGLTQTCETLKLQLPNCIEDPVPIVLRNF	736						
Qy	721	SLVGTPLSAFGLNRPVLAEDAQLFTALPPEKNCNDNICODDLSITFSFMSLCLVVG	780						
Db	737	SLVGTPLSAFGLNRPVLAEDAQLFTALPPEKNCNDNICODDLSITFSFMSLCLVVG	796						
Qy	781	GPREFNVTVVRNMGDSYRTQVTFPPLDLSYRKVSTIQLQNSQSRWLAACESASTEV	840						
Db	797	GPREFNVTVVRNMGDSYRTQVTFPPLDLSYRKVSTIQLQNSQSRWLAACESASTEV	856						
Qy	841	SGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	900						
Db	857	SGALKSTCSINHPIIPENSEVTFNITFDVDSKASLGNKLLKANVTSENMPRTNKTEF	916						
Qy	901	QLELPVKYAVYVMTVSHGVSTKYLNTFTASNTSRVMQHQYQVSNLQORSILPISLVFLPV	960						
; RESULT 9									
US-09-688-307A-3									
; Sequence 3, Application US/09688307A									
; Patent No. 6432404									
; GENERAL INFORMATION:									
; APPLICANT: Gallatin, Michael W.									
; APPLICANT: Van der Vieren, Monica									
; TITLE OF INVENTION: No. 6432404el Human Beta-2									
; FILE REFERENCE: 27866/36646									
; CURRENT APPLICATION NUMBER: US/09/688,307A									
; CURRENT FILING DATE: 2000-10-13									
; PRIOR APPLICATION NUMBER: 09/193,043									
; PRIOR FILING DATE: 1998-11-16									
; PRIOR APPLICATION NUMBER: 08/605,672									
; PRIOR FILING DATE: 1996-02-22									
; PRIOR APPLICATION NUMBER: 08/173,497									
; PRIOR FILING DATE: 1993-12-23									
; PRIOR APPLICATION NUMBER: 08/286,889									
; PRIOR FILING DATE: 1994-08-05									
; PRIOR APPLICATION NUMBER: 08/362,652									
; PRIOR FILING DATE: 1994-12-21									
; PRIOR APPLICATION NUMBER: 08/943,363									
; PRIOR FILING DATE: 1997-10-03									
; NUMBER OF SEQ ID NOS: 114									
; SOFTWARE: Patent In Ver. 2.0									
; SEQ ID NO 3									
; LENGTH: 1153									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-09-688-307A-3									
Query Match 99.9%; Score 5868; DB 4; Length 1153;									
Best Local Similarity 99.6%; Pred. No. 0;									
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;									
Qy	1	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI	60						
Db	17	FNLDTENAMTFOENARGFGQSVVQLQGSRRVVVGAPQEIIVAAHQSGSLYQCDYSTGSCBPI	76						
Qy	61	RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLFGSNLRQOPQK	120						
Db	77	RLQVPEAVNMSLGLSLAATTSPPOLLACGPTVHQTCSNTYVKGCLFGLFGSNLRQOPQK	136						
Qy	121	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSEEF	180						
Db	137	FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRMKEFVSTVMEQKKSKTLFSLMOYSEEF	196						
Qy	181	RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIRELLNITNGARKNAFKILIVI	240						
Db	197	RIHFTFKFQNNPNRSLVKPITQLGRTHATGVRKVIRELLNITNGARKNAFKILIVI	256						
Qy	241	TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN	300						
Db	257	TDGKFGDPLGYEDVPEADREGVIRYVIGVDAPFRSEKSRQELNTIASKPPRDHVFQVN	316						
Qy	301	NFEALKTIONOLREKIFAIEGTQTGSSSPHEHMSQEGFSAATISNGPLLSITVGSYDWAG	360						

```
Db 317 NFEALKTIONLREKIFAIEGTQTGSSSSFEHMSQEGFAAITSNGPILLSVTGSDYDAG 376
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVGLGAPYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVGLGAPYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVYBQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVYBQTRGGQVSVCP 496
Qy 481 PRGORARWOCDAVLYGEGQGPWGRFGAALTIVLGVNGDKLTDAVIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLYGEGQGPWGRFGAALTIVLGVNGDKLTDAVIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGTOTCETLKLQPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRTRQTVLGTOTCETLKLQPNCIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFNGNLRPVLAEADAQRLFTALFPFEKNCNGNDNICODDLSITFSFMSLDC 780
Db 737 SLVGTPLSAFNGNLRPVLAEADAQRLFTALFPFEKNCNGNDNICODDLSITFSFMSLDC 796
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLRACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLRACESASSTEV 856
Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 900
Db 857 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 916
Qy 901 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSIPLSLFLVPV 960
Db 917 QLELPVKYAVYVMTSHGVSTKYLNFTASNTSRVMQHOYQVSNLQORSIPLSLFLVPV 976
Qy 961 RLMQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1020
Db 977 RLMQTVIMDRPQVTFSENLSSTCHTKERLPSHSDFLAELRKAPVNCSTAVCQRIQCDIP 1036
Qy 1021 FFGIQEEFNATLKGNSLPDWYIKTSHNHLIVSTABILFNDVSFTLLPQGFVRSQDET 1080
Db 1037 FFGIQEEFNATLKGNSLPDWYIKTSHNHLIVSTABILFNDVSFTLLPQGFVRSQDET 1096
Qy 1081 KVEPFEPVNPPLIVGSSVGGILLLALITAALYKLGFEKRYKQKMMSEGGPPGAEPO 1137
Db 1097 KVEPFEPVNPPLIVGSSVGGILLLALITAALYKLGFEKRYKQKMMSEGGPPGAEPO 1153

RESULT 10
US-09-350-259-3
; Sequence 3, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; EARLIER FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
```

```
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1153
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-350-259-3
```

```
Query Match 99.9%; Score 5868; DB 4; Length 1153;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1133; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 FNLDTENAMTFQENARFGQSVVLOQSGRVVVGAPQEIIVAAORGSLVYCDYSTGSCPEI 60
Db 17 FNLDTENAMTFQENARFGQSVVLOQSGRVVVGAPQEIIVAAORGSLVYCDYSTGSCPEI 76
Qy 61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQ 120
Db 77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVVKGLCFLFGSNLRQOPQ 136
Qy 121 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQKKSKTLFSLMQYSEEF 180
Db 137 FPALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFVSTVMEQKKSKTLFSLMQYSEEF 196
Qy 181 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGVRKVIPELLNITNGARKNAFKILIVI 240
Db 197 RIHFTFKFQNNPNRSLVKPITQLLGRTHATGIRKVKRELFNITNGARKNAFKILVVI 256
Qy 241 TDGEKFGDPLGYEDVPIEADREGVIRYVIGVGDAPFSEKSRQBELNTIASKPPRDHVFQVN 300
Db 257 TDGEKFGDPLGYEDVPIEADREGVIRYVIGVGDAPFSEKSRQBELNTIASKPPRDHVFQVN 316
Qy 301 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFAAITSNGPILLSVTGSDYDAG 360
Db 317 NFEALKTIQNLREKIFAIEGTQTGSSSSFEHMSQEGFAAITSNGPILLSVTGSDYDAG 376
Qy 361 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVGLGAPYOHIGLVAMFR 420
Db 377 GVFLYTSKEKSTFINMTVRDSDMDNDAYLGAAAAIILNRVOSLVGLGAPYOHIGLVAMFR 436
Qy 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVYBQTRGGQVSVCP 480
Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYVYBQTRGGQVSVCP 496
Qy 481 PRGORARWOCDAVLYGEGQGPWGRFGAALTIVLGVNGDKLTDAVIGAPGEEDNRGAVYLF 540
Db 497 PRGORARWOCDAVLYGEGQGPWGRFGAALTIVLGVNGDKLTDAVIGAPGEEDNRGAVYLF 556
Qy 541 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 600
Db 557 HGTSGSGISPSHSQRIAGSKLSPLQYFGQSLSGQDLTMDGLVDLTVGAQGHVLLRSQ 616
Qy 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 660
Db 617 PVLRVKAIMFNPREVARNVFECDQVVKGEAGEVRVCLHVQKSTRDRLRREGQIQSVVT 676
Qy 661 YDLALDSGRPHSRVAFNETKNSRTRQTVLGTOTCETLKLQPNCIEDPVSPVILRLNF 720
Db 677 YDLALDSGRPHSRVAFNETKNSRTRQTVLGTOTCETLKLQPNCIEDPVSPVILRLNF 736
Qy 721 SLVGTPLSAFNGNLRPVLAEADAQRLFTALFPFEKNCNGNDNICODDLSITFSFMSLDC 780
Db 737 SLVGTPLSAFNGNLRPVLAEADAQRLFTALFPFEKNCNGNDNICODDLSITFSFMSLDC 796
Qy 781 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLRACESASSTEV 840
Db 797 GPREFNVTVVRNDEGDSYRTQVTFPPFLDLSYRKVSTLQNRQSQRSLRACESASSTEV 856
Qy 841 SGALKSTSCSINHPIIPENSEVTFNITFDVDSKASLGKLLKANVTSENNMPRTNTEF 900
```

857 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 916  
901 QLELPVKYAVYVMTSHGVSSTKYLNFTASNTSRVMOHQYQVSNLQORSILPISLVFLVPV 960  
917 QLELPVKYAVYVMTSHGVSSTKYLNFTASNTSRVMOHQYQVSNLQORSILPISLVFLVPV 976  
961 RLNQTVWDRPQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORIQCDIP 1020  
977 RLNQTVWDRPQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORIQCDIP 1036  
1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVRSTET 1080  
1037 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVRSTET 1096  
1081 KVPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1137  
1097 KVPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1153

RESULT 11

US-08-476-062A-43  
Sequence 43, Application US/08476062A  
Patent No. 5877275  
GENERAL INFORMATION:  
APPLICANT: Arnsout, M. Amin  
TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY  
TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: US  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,062A  
FILING DATE: 07-JUN-1995  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 08/216,081  
FILING DATE: 21-MAR-1994  
APPLICATION NUMBER: 07/637,830  
FILING DATE: 04-JAN-1991  
APPLICATION NUMBER: 07/539,842  
FILING DATE: 18-JUN-1990  
APPLICATION NUMBER: 07/212,573  
FILING DATE: 28-JUN-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00786/068003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1152 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-476-062A-43

Query Match 99.4%; Score 5837.5; DB 2; Length 1152;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

1 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAANORGLYQCDYSTGSCPEI 60  
17 FNLDTENAMTFQENARFGQSVVQLQGSRRVVVGAPOEIVAAANORGLYQCDYSTGSCPEI 76  
61 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 120  
77 RLQVPVEAVNMSLGLSLAATTSPQLLACGPTVHQTCSNTYVKGICFLFGSNLRQOPQK 136  
121 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEFVSTVMEQLKSKTILFSLMQYSEEF 180  
137 FPEALRGCPQEDSDIAFLIDGSGSIIPHDFRRMKFEFVSTVMEQLKSKTILFSLMQYSEEF 196  
181 RIHFTFKFQNNPNPSLVKPIITQLLGRTHATATGVRKVIKRELLININGAKNAFKILJVI 240  
197 RIHFTFKFQNNPNPSLVKPIITQLLGRTHATATGVRKVIKRELLININGAKNAFKILJVI 256  
241 TDGEKFGDPLGYEDVPEADREGVIRVYGVDAFSEKSRQELINTIASKPPDRHVFQVN 300  
257 TDGEKFGDPLGYEDVPEADREGVIRVYGVDAFSEKSRQELINTIASKPPDRHVFQVN 316  
301 NFEALKTIONQLREKIFAIEGTOTGSSSSPEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 360  
317 NFEALKTIONQLREKIFAIEGTOTGSSSSPEHEMSQEGFSAAITSNGLPLSTVGSYDWAG 376  
361 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLGAPRYQHIGLVAMER 420  
377 GVFLYTSKEKSTFINMTRVDSMDNDAYLGYAAAILLRNVQSLVGLGAPRYQHIGLVAMER 436  
421 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQSVCP 480  
437 QNTGMWESNANVKGTOIGAYFGASLCSVDVDSNGSTDLVLIGAPHYVEQTRGGQSVCP 496  
481 PRGORARWQCDVLYGEOGQPGWRFGAALTIVLGVNVDGDKLTDVAIGAPEGEDNRGAVLYF 540  
497 PRG-RARWQCDVLYGEOGQPGWRFGAALTIVLGVNVDGDKLTDVAIGAPEGEDNRGAVLYF 555  
541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGCHVILLRSQ 600  
556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGQDLTMDGLVDLTGAGCHVILLRSQ 615  
601 PVLRVKAIMENPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDLREGOIQSVVT 660  
616 PVLRVKAIMENPREVARNVFECDQVVKKEAGEVRVCLHVOKSTRDLREGOIQSVVT 675  
661 YDLALDSGRPHSRVAFNETKSTRTOVLGLTQTCETLKLQLPNCIEDVPVIVLRNLF 720  
676 YDLALDSGRPHSRVAFNETKSTRTOVLGLTQTCETLKLQLPNCIEDVPVIVLRNLF 735  
721 SLVGTPLSAFONLRPVLAEDAQRLLFTALFPPEKNCNDNIQDDLSITFSFMSLCLVVG 780  
736 SLVGTPLSAFONLRPVLAEDAQRLLFTALFPPEKNCNDNIQDDLSITFSFMSLCLVVG 795  
781 GPREFNVTVVRNDEDSYRTQVTFEPLDLSYRKVSTLQNRORSWRLACESASSTEV 840  
796 GPRESNVTVVRNDEDSYRTQVTFEPLDLSYRKVSTLQNRORSWRLACESASSTEV 855  
841 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 900  
856 SGALKSTSCSINHPIFPENSEVTNITFDVDSKASLGNKLLKANVTSENNPRTNKTFF 915  
901 QLELPVKYAVYVMTSHGVSSTKYLNFTASNTSRVMOHQYQVSNLQORSILPISLVFLVPV 960  
916 QLELPVKYAVYVMTSHGVSSTKYLNFTASNTSRVMOHQYQVSNLQORSILPISLVFLVPV 975  
961 RLNQTVWDRPQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORIQCDIP 1020  
976 RLNQTVWDRPQVTFSENLSTCHTKERLPSSHDFLAELRKAPVNVCSIAVCORIQCDIP 1035  
1021 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVRSTET 1080  
1036 FFGIQEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVRSTET 1095  
1081 KVPEFVFNPLPLIVGSSVGGLLLLALITAAALYKLGFFKQYKQMMSEGGPPGAEPQ 1137

Db 1096 KVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1152

## RESULT 12

PCT-US96-01314-43

; Sequence 43, Application PC/TUS9601314

; GENERAL INFORMATION:

; APPLICANT: M. Amin Arnaout

; TITLE OF INVENTION: METHODS FOR IDENTIFYING INTEGRIN

; TITLE OF INVENTION: ANTAGONISTS

; NUMBER OF SEQUENCES: 78

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish &amp; Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: Massachusetts

; COUNTRY: U.S.A.

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

; COMPUTER: IBM PS/2 Model 50Z or 55SX

; OPERATING SYSTEM: MS-DOS (Version 5.0)

; SOFTWARE: WordPerfect (Version 5.1)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/01314

; FILING DATE: 30-JAN-96

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/380,167

; FILING DATE: 30-JAN-95

; ATTORNEY/AGENT INFORMATION:

; NAME: John W. Freeman

; REGISTRATION NUMBER: 29,066

; REFERENCE/DOCKET NUMBER: 00786/267001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 542-5070

; TELEFAX: (617) 542-8906

; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 43:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1152

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: linear

PCT-US96-01314-43

## Query Match

Best Local Similarity 99.4%; Score 5837.5; DB 5; Length 1152;

Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

QY 1 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANORGLYQCDYSTGSCBPI 60

Db 17 FNLDTENAMTFQENARGFGQSVVQLQGSRRVVVGAPQEIIVAAANORGLYQCDYSTGSCBPI 76

QY 61 RLQVPVEAVNMSLGLSLAATTSPPQLLAGPTVHTQCSNTYVKGCLFLFGSNLRQOQPK 120

Db 77 RLQVPVEAVNMSLGLSLAATTSPPQLLAGPTVHTQCSNTYVKGCLFLFGSNLRQOQPK 136

QY 121 PPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKBFVSTVMQKSKTLFSLMQYSEEF 180

Db 137 PPEALRGCPQSDSDIAFLIDGSGSIIPHDPRMKBFVSTVMQKSKTLFSLMQYSEEF 196

QY 181 RIHFTFKFQNNPNRSLVKPTQLLGRTHATGVRKVIKRELLNTNGARKNAFKILVI 240

Db 197 RIHFTFKFQNNPNRSLVKPTQLLGRTHATGVRKVIKRELLNTNGARKNAFKILVI 256

QY 241 TDGEXFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDHVFQV 300

Db 257 TDGEXFGDPLGYEDVPEADREGVIRYVIGVDGAFRSEKSRQELNTIASKPPRDHVFQV 316

QY 301 NFEALKTKQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITNGPLLTGVSVDWAG 360

Db 317 NFEALKTKQNLREKIFAIEGTQSGSSSFEHMSQEGFSAAITNGPLLTGVSVDWAG 376

QY 361 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAALILRNVRQSVLVLGAPRYOHIGLVAMPR 420

Db 377 GVFLYTSKEKSTPINNTRVDSMDNDAYLGVAALILRNVRQSVLVLGAPRYOHIGLVAMPR 436

QY 421 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQGVSCPL 480

Db 437 QNTGMWESNANVKGTQIGAYFGASLCSVDVDSNGSTDVLVIGAPHYYEOTRGQGVSCPL 496

QY 481 PRQRRARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 540

Db 497 PRG-RARWQCDVLYGEGQGPWGRFGAALTVLGDVNGDKLTDVAIGAPGEDNRGAYVLF 555

QY 541 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGGDLTMDGLVDLTGVAQGHVLLRSQ 600

Db 556 HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLSGGGDLTMDGLVDLTGVAQGHVLLRSQ 615

QY 601 PVLRVKAIMFNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLREGIOISVVT 660

Db 616 PVLRVKAIMFNPREVARNVFECDQVVKGEAGVRVCLHVOKSTRDLREGIOISVVT 675

QY 661 YDLALDSGRPHSRVAVNETKSTRQTQVLGTLTQCTETLKLQLPNCIEDPVPSPVILRLNF 720

Db 676 YDLALDSGRPHSRVAVNETKSTRQTQVLGTLTQCTETLKLQLPNCIEDPVPSPVILRLNF 735

QY 721 SLVGTPLSAFNGNLRPVLAEDAORLFTALPFFPKNCNDNICQDDLSITFSFMSLDCLVVG 780

Db 736 SLVGTPLSAFNGNLRPVLAEDAORLFTALPFFPKNCNDNICQDDLSITFSFMSLDCLVVG 795

QY 781 GPREFNVTVTVRNDGEDSVRTQVTFPFDLDSYRKVSTLONORSQBSWELACESASSTEV 840

Db 796 GPREFNVTVTVRNDGEDSVRTQVTFPFDLDSYRKVSTLONORSQBSWELACESASSTEV 855

QY 841 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLLKANTVSENMPRTNKTFF 900

Db 856 SGALKSTSCSINHPIFPENSEVTFNITFDVDSKASLGNKLLKLLKANTVSENMPRTNKTFF 915

QY 901 QLELPVKYAVYVWVTSYHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLPISLVLVVPV 960

Db 916 QLELPVKYAVYVWVTSYHGVSTKYLNFTASENTSRVMQHOYQVSNLQGRSLPISLVLVVPV 975

QY 961 RLNQTVIWDROPVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNCSTAVCQRCQCDIP 1020

Db 976 RLNQTVIWDROPVTFSENLSSTCHTKERLPSSHSDFLAELRKAPVNCSTAVCQRCQCDIP 1035

QY 1021 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPQCGAFVRSQTF 1080

Db 1036 PFGIQEEFNATLKGNSLFDWYIKTSHNHLIIVSTAELFNDSVFTLLPQCGAFVRSQTF 1095

QY 1081 KVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1137

Db 1096 KVEPFEVNPPLIVGSSVGLLLALITAAALYKLGFFKQYKDMSEGGPPGABPQ 1152

## RESULT 13

5424399-2

; Patent No. 5424399

; APPLICANT: ARNAOUT, M. AMIN

; TITLE OF INVENTION: HUMAN CR3a/b HETERODIMERS

; NUMBER OF SEQUENCES: 12

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/78,871

; FILING DATE: 16-JUN-1993

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 539,842

; FILING DATE: 18-JUN-1990

; APPLICATION NUMBER: 212,573

; FILING DATE: 28-JUN-1988

; SEQ ID NO: 2

; LENGTH: 1152

5424399-2

Query Match

99.4%; Score 5837.5; DB 6; Length 1152;

Best Local Similarity 99.4%; Pred. No. 0;									
Matches 1130; Conservative 3; Mismatches 3; Indels 1; Gaps 1;									
Qy	1	FNLDTENAMTFQENARGFGQSVVQLOQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI	60						
Db	17	FNLDTENAMTFQENARGFGQSVVQLOQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI	76						
Qy	61	RLQVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSLNRQPOK	120						
Db	77	RLQVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSLNRQPOK	136						
Qy	121	FPFALRGCPQEDSDIAFLIDGSGIIPHDPRRKEFVSTWEOQKSKTILFSLMQYSEEF	180						
Db	137	FPFALRGCPQEDSDIAFLIDGSGIIPHDPRRKEFVSTWEOQKSKTILFSLMQYSEEF	196						
Qy	181	RIHFTKFEQNNPNRSLVKPIITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI	240						
Db	197	RIHFTKFEQNNPNRSLVKPIITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI	256						
Qy	241	TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSOELNTIASKPRDHVFOVN	300						
Db	257	TDGEKFGDPLGYEDVPEADREGVIRYVIGVDAFRSEKSOELNTIASKPRDHVFOVN	316						
Qy	301	NFALKTIQNLREKIPAIETGTOTGSSSPHEHMQSGFSAATISNGPLLSTVGSYDWAG	360						
Db	317	NFALKTIQNLREKIPAIETGTOTGSSSPHEHMQSGFSAATISNGPLLSTVGSYDWAG	376						
Qy	361	GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNVQSVLGPAPRYOHLGVAMFR	420						
Db	377	GVFLYTSKESKSTFINMTRVDSMDNDAYLGAAAILLRNVQSVLGPAPRYOHLGVAMFR	436						
Qy	421	QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSVCP	480						
Db	437	QNTGMWESNANVKTQIGAYFGASLCSVDVDSNGSTDVLIGAPHYYEOTRGGOVSVCP	496						
Qy	481	PRGORARWQCDVILYGGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYL	540						
Db	497	PRG-RARWQCDVILYGGQGPWGRFGAALTVLGVDNGDKLTDVAIGAPGEDNRGAYL	555						
Qy	541	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ	600						
Db	556	HGTSGSGISPSHSQRIAGSKLSPRLQYFGQSLGGQDLTMDGLVDLTVGAQGHVLLRSQ	615						
Qy	601	PVLRYKAIMENPREVARNVFECDQVVKGEAGEVVCVCHVOKSTDRRREGIOQSVVT	660						
Db	616	PVLRYKAIMENPREVARNVFECDQVVKGEAGEVVCVCHVOKSTDRRREGIOQSVVT	675						
Qy	661	YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLENCIEDPVSPIVLRNF	720						
Db	676	YDLALDSGRPHSRVAFNETKNSRRQTQVLGLTQTCETLKLQLENCIEDPVSPIVLRNF	735						
Qy	721	SLVGTPLSAFONLRPVLAEDAQRFTALFPPEKNCMDNICQDDLSITFSFMSLDCLVWG	780						
Db	736	SLVGTPLSAFONLRPVLAEDAQRFTALFPPEKNCMDNICQDDLSITFSFMSLDCLVWG	795						
Qy	781	GPREFNVTVVRNDGDSYRQVTFPPFLDLSYRKVSTLQNRORSORSLRACESASSTEV	840						
Db	796	GPREFNVTVVRNDGDSYRQVTFPPFLDLSYRKVSTLQNRORSORSLRACESASSTEV	855						
Qy	841	SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF	900						
Db	856	SGALKSTSCSINHPIFFENSEVTNITFDVDSKASLGNKLLKANVTSENMPRTNKTFF	915						
Qy	901	QLELPVKYAVNMVTSHGVTSKYLNFTASENTSRVMQHOYQVSNLQORSLPISLFLVPV	960						
Db	916	QLELPVKYAVNMVTSHGVTSKYLNFTASENTSRVMQHOYQVSNLQORSLPISLFLVPV	975						
Qy	961	RLNQTVIWRDQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1020						
Db	976	RLNQTVIWRDQVTFSENLSTCHTKERLPSHSDFLAELRKAPVNVCSIAVCQRIQCDIP	1035						
Qy	1021	FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVRQOTET	1080						

Db	1036	FFGIOEEFNATLKGNSLFDWYIKTSHNHLIVSTAEILFNDVSFTLLPGQGAFAVRQOTET	1095						
Qy	1081	KVEPEFVNPPLPIVVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ	1137						
Db	1096	KVEPEFVNPPLPIVVGSSVGGLLLLALITAAALYKLGFFKQYKDMMSGEGPPGAEPQ	1152						
RESULT 14									
US-08-476-062A-44									
; Sequence 44, Application US/08476062A									
; Patent No. 5877275									
; GENERAL INFORMATION:									
; APPLICANT: Arnaout, M. Amin									
; TITLE OF INVENTION: CONTROLLING CELLULAR IMMUNE/INFLAMMATORY									
; TITLE OF INVENTION: RESPONSES WITH BETA2 INTEGRINS									
; NUMBER OF SEQUENCES: 53									
; CORRESPONDENCE ADDRESS:									
; ADDRESSEE: Fish & Richardson P.C.									
; STREET: 225 Franklin Street									
; CITY: Boston									
; STATE: MA									
; COUNTRY: US									
; ZIP: 02110-2804									
; COMPUTER READABLE FORM:									
; MEDIUM TYPE: Diskette									
; COMPUTER: IBM Compatible									
; OPERATING SYSTEM: Windows95									
; SOFTWARE: FastSeq for Windows Version 2.0									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: US/08/476,062A									
; FILING DATE: 07-JUN-1995									
; PRIOR APPLICATION DATA:									
; APPLICATION NUMBER: 08/216,081									
; FILING DATE: 21-MAR-1994									
; APPLICATION NUMBER: 07/637,830									
; FILING DATE: 04-JAN-1991									
; APPLICATION NUMBER: 07/539,842									
; FILING DATE: 18-JUN-1990									
; APPLICATION NUMBER: 07/212,573									
; FILING DATE: 28-JUN-1988									
; ATTORNEY/AGENT INFORMATION:									
; NAME: Freeman, John W.									
; REGISTRATION NUMBER: 29,066									
; REFERENCE/DOCKET NUMBER: 00786/068003									
; TELECOMMUNICATION INFORMATION:									
; TELEPHONE: 617/542-5070									
; TELEFAX: 617/542-8906									
; TELEX: 200154									
; INFORMATION FOR SEQ ID NO: 44:									
; SEQUENCE CHARACTERISTICS:									
; LENGTH: 1163 amino acids									
; TYPE: amino acid									
; TOPOLOGY: linear									
; MOLECULE TYPE: protein									
US-08-476-062A-44									
Query Match 59.1%; Score 3473; DB 2; Length 1163;									
Best Local Similarity 61.3%; Pred. No. 4, 2e-285;									
Matches 692; Conservative 138; Mismatches 293; Indels 6; Gaps 4;									
Qy	1	FNLDTENAMTFQENARGFGQSVVQLOQSGRVVVGAPQEIIVAAQNGSLYQCDYSTGSCPEI	60						
Db	20	FNLDTBELTAFRVDSAGFGDSVWQYANSVWVGAPQKITAANQTGGLYQCCYSTGACPEI	79						
Qy	61	RLQVPVAVNMVSLGLSLAATTSPPOLLACGPTVHOTCSENTYVKGICFLFGSLNRQPOK	120						
Db	80	GLQVPPFAVNMVSLGLSLAATTSPPOLLACGPTVHHECGRNMYLTGLCFLLGPT--QLTQR	137						
Qy	121	FPFALRGCPQEDSDIAFLIDGSGIIPHDPRRKEFVSTWEOQKSKTILFSLMQYSEEF	180						
Db	138	LPVSRQCPQEQDILFIDGSGISSENFAFMFVRAVISQFQRPSTQSLQFSNKF	197						
Qy	181	RIHFTKFEQNNPNRSLVKPITQLLGRTHATGVRKVIRELLNITNGARKNAFKILIVI	240						





541	QY	HTSGSGISPSHSQRIAGSKLSPRIQYFGQSLSGGQDLTMDGLVLTVGAQGHVLLRSQ	600
557	Db	HCVLGFSPISPSHSQRIAGSQLSRRLQYFGQALSGGQDLTQDGLVDLAVGARQVLLRLTR	616
601	QY	PVLRVKAIWEPNPREVARNVPCNDQVYKGEAVRVCLHVOKSTRODLREGQIQSVVT	660
617	Db	PVLVGVSMQFIPAEIPRSAFECEBQVSEQTLVQSNICLYIDKRSKNLLGSRDIQSSVT	676
661	QY	YDLALDGRPHGRAVFNKSTRRQTQVLGTLQTCTTKLQLPNCIEDPSPIVLRLNF	720
677	Db	LDLALDPGLSPRAFPQETKNBSLSEVRVGLKACHCENFNLLPSCVEDSVTPITLRLNF	736
721	QY	SLVGTPLSPAGNLRPVLAEADAQRLFTALPFPFKNCGNDNICODDLSITPFSMSLCIVVG	780
737	Db	TLVGKPLLAFRNLRPMLAALAQRYFTASLPFFKNCGADHICODNLGISFPFGLKSLVG	796
781	QY	GRPEFNVTYVRNDGSDSYRTQVTFPPLDLSYRKVSTLQNRQSRSWRLACESASSTEV	840
797	Db	SNLELNAEVMWMDGSDSYGTTITFSPHAGLSYRYVAEGKQKQGLRSJHLTCDSPVGV--	854
841	QY	SGALKSTCSINHPFIPENSEVFNTITFQVDSKASLGNKLLKANVTSNNPRTNKTFF	900
855	Db	SGQTWSTSCRNHLIFRGGAQITFLATFQVSPKAVLGRILLTANVSSNNPRTSKTTF	914
901	QY	QLELPVKYAVYVMTVSHGVSTKYLVNTAS--ENTSRVMQHQYQVSNLQBSLPLSVFLVLP	959
915	Db	QLELPVKYAVYTVVSSHEQTKYLVNFSESEKESHVAMHRYQVNNLQORDLPSVINFVWP	974
960	QY	VLRLNTOIVMDRPOVTFSENLSSTCHTKERLPUSDHFLAELRKAPVNVCSITAVQRTQCDDI	1019
975	Db	VELNQEAVMMDVEVSHQNPFSKCSSEKIAPPASDFLAHQKNPVLDCSIAGCLRFECDV	1034
1020	QY	PFFGIOEEFNATLKGNSLFDWYIATXSHNHLIIVSTABILFNDSVFTLLPQGFVRSQTE	1079
1035	Db	PSFSVQEELDFTLKGNSLPCWVRQILQKXSVSVVAEITPDTSVYSOLPQGEAFMEAQTT	1094
1080	QY	TKVPEPEVPNPLPIVGVSSVGGLLLLLAIITAAIYKLGFFKRYKDWME	1128
1095	Db	TVLEKYVHNPTPLIVGSSITVGGLLLLLAIITAVIYKGVFFKRYKEMEE	1143

Search completed: November 9, 2004, 12:21:14

Job time : 28 secs